

Reconstrução Tridimensional de uma Subestação de Energia Elétrica para Localização de Equipamentos a Partir de Termogramas

Francisco I. S. Ganacim¹

Tecgraf/PUC-Rio, Rio de Janeiro, RJ

Leonardo G. Fernandes² Eduardo Schumacher³ Nara Bobko⁴ Eduardo F. R. Romaneli⁵
UTFPR, Curitiba, PR

Resumo. Neste trabalho, propomos e implementamos uma forma de realizar a associação entre imagens termográficas de equipamentos elétricos com a localização destes equipamentos na subestação de energia elétrica. Para tal, realizamos a reconstrução tridimensional da subestação de energia utilizando algoritmos de *Structure from Motion* e correspondência de pontos-chave. Após concluir a reconstrução tridimensional, comparamos os equipamentos presentes na imagem termográfica com os da reconstrução por meio da correspondência de pontos-chave. Isso permite uma associação eficiente entre as imagens termográficas e a localização dos respectivos equipamentos na subestação de energia.

Palavras-chave. Inspeções termográficas Automatizadas, Subestações de Energia, Localização de Equipamentos, Reconstrução Tridimensional, Structure from Motion.

1 Introdução

Inspeções termográficas em Subestações de Energia Elétrica (SE) possuem um papel importante na manutenção preditiva por possibilitarem a identificação de equipamentos com temperaturas anormais e, assim, detectar precocemente falhas nos equipamentos elétricos [5]. Em geral, tanto as inspeções como as análises dos termogramas são realizadas de forma manual. A fim de viabilizar uma manutenção preditiva contínua, a automatização deste processo é fundamental, permitindo inspeções frequentes e reduzindo o tempo de análise dos termogramas e dos custos operacionais.

O Projeto *Desenvolvimento de Metodologia para Análise Automática de Imagens Térmicas*⁶, do qual este trabalho faz parte, propõe uma forma de realizar este processo de automatização. Neste projeto o processo de automatização é dividida nas seguintes etapas: captação e processamento das imagens térmicas, identificação dos equipamentos elétricos presentes nestas imagens, detecção de temperaturas anormais nas regiões dos termogramas correspondentes aos equipamentos, e localização destes equipamentos na SE. A última etapa, que é o foco deste trabalho, é essencial para que a equipe técnica saiba exatamente qual o equipamento requer manutenção. Neste contexto, o objetivo deste trabalho é realizar a associação entre os termogramas (imagens bidimensionais) e a localização precisa desses equipamentos na SE (posição tridimensional). Essa associação torna-se

¹ganacim@gmail.com

²leonardofernandes@alunos.utfpr.edu.br

³schumacher@alunos.utfpr.edu.br

⁴narabobko@utfpr.edu.br

⁵felix@utfpr.edu.br

⁶Projeto PD-2866- 0528/2020 - Desenvolvimento de Metodologia para Análise Automática de Imagens Térmicas. Projeto P&D executado pela COPEL-DIS e regulado pela ANEEL.

mais desafiadora devido ao fato de uma SE possuir vários equipamentos do mesmo tipo (por exemplo, várias chaves-seccionadoras). Desta forma não basta identificar o tipo de objeto, precisamos identificar sua posição específica com base apenas na imagem presente no termograma.

Para tal, realizamos uma reconstrução tridimensional (3D) da SE a partir de um conjunto sequencial de imagens da SE utilizando uma técnica de *Structure from Motion* (SfM) [7]. Esta reconstrução irá gerar um conjunto de pontos que representa todos os objetos e estruturas da SE (Seção 2). Em seguida, identificamos os equipamentos presentes na SE dentro desta reconstrução 3D utilizando uma envoltória convexa do equipamento utilizando-se múltiplos pontos de vista. Isto é, obtemos quais pontos da reconstrução correspondem a cada um dos equipamentos da SE a serem monitorados, o que nos permite identificar a posição destes equipamentos na SE (Seção 3). Por fim, realizamos a associação dos equipamentos detectados nos termogramas (provenientes das inspeções termográficas) aos equipamentos pré-identificados na reconstrução 3D, permitindo assim localizar precisamente o equipamento de interesse (Seção 4).

A reconstrução 3D adotada neste estudo foi baseada no artigo de Moulon et al. [8], pois, assim como no nosso caso, não requer o uso de sensores adicionais (tais como sensores de profundidade). No referido artigo, os autores buscam obter uma reconstrução 3D RGB a partir de termogramas, resultando em uma nuvem de pontos térmicos. Isto é, os pontos representam os objetos da cena e a cor dos pontos a temperatura. Essa abordagem permite explorar simultaneamente os aspectos geométricos (posição dos objetos no espaço) e térmicos. No entanto, para lidar com o problema, é necessária uma reconstrução densa, a fim de que partes dos objetos com temperaturas elevadas não sejam descartados. No nosso caso, não estamos interessados em analisar termicamente toda a cena, mas sim objetos individuais e em diferentes momentos, pois cada inspeção pode apresentar variações nas temperaturas dos equipamentos. É importante ressaltar que consideramos a possibilidade de inspeções contínuas, capturando imagens térmicas em intervalos curtos de tempo (utilizando um AGV, por exemplo). Com base nisso, optamos por realizar uma reconstrução 3D da SE de forma esparsa, levando em conta apenas os aspectos geométricos. Em seguida, associamos os termogramas das inspeções aos equipamentos presentes nessa reconstrução 3D.

2 Reconstrução tridimensional

A reconstrução tridimensional de uma cena é uma estrutura de dados que representa a informação geométrica de um dado ambiente e dos objetos que ela contém. Concretamente, a reconstrução é representada como uma nuvem não estruturada de pontos tridimensionais, sendo que cada ponto é relacionado com, ao menos, três das imagens usadas para a reconstrução. Para a reconstrução 3D da SE, utilizamos uma família de algoritmos conhecidos como *Structure from Motion* (SfM). Mais especificamente, utilizamos um método incremental de SfM baseado no algoritmo apresentado por Schonberger e Frahm [10].

O ponto de partida do processo foi um vídeo da SE realizado por uma câmera comum (definição *fullhd* a 120fps). Em seguida, realizamos uma amostragem deste vídeo obtendo então um conjunto de imagens sequenciais da SE. Optamos por uma amostragem adaptativa a fim de reduzir a quantidade total de imagens e, ao mesmo tempo, evitar perder informações importantes (detalhamento adiante). A Figura 1 apresenta exemplos de *frames* utilizados para a reconstrução.



Figura 1: Exemplo de *frames* utilizados para a reconstrução. Fonte: Autor (2022).

Em seguida, identificamos os pontos-chaves de cada uma das imagens. Os pontos-chaves de uma imagem são os pontos mais relevantes para compreender a estrutura e os objetos (equipamentos) presentes em uma imagem (invariantes por pequenas translações ou movimentos de câmera). Para tal, utilizamos o algoritmo *Scale-invariant feature transform* (SIFT) [6] que, em linha gerais, encontra estes pontos e atribui a cada um deles um descritor que captura as características geométricas locais.

O próximo passo será, para cada par de imagens, encontrar a matriz essencial que fornece a geometria da transição de pontos de vista. Para isto precisamos fazer uma correspondência (um pareamento) entre os pontos-chaves em imagens diferentes. Isto é, encontrar pontos-chaves de duas imagens sequenciais que possuam o mesmo descritor SIFT (mais precisamente, a distância entre os descritores SIFT dos dois pontos deverá ser suficientemente pequena). Todavia esta forma de fazer a correspondências entre os pontos-chaves acaba gerando várias correspondências incorretas (tanto falso-positivo quanto falso-negativo). A fim de remover os falsos-positivos, realizamos uma filtragem nos pontos pareados utilizando o algoritmo probabilístico *Random Sample Consensus* (RANSAC) [3]. Obtemos então um conjunto de pares de pontos-chaves entre as duas imagens e matriz essencial destas duas imagens.

A Figura 2 mostra um par de *frames* usados na reconstrução que observam a mesma região da SE mas a partir de pontos de vista distintos. Na parte inferior desta figura podemos ver em vermelho todos os pontos-chaves detectados e, em linhas verdes, os pares de pontos cujas correspondências foram identificados.

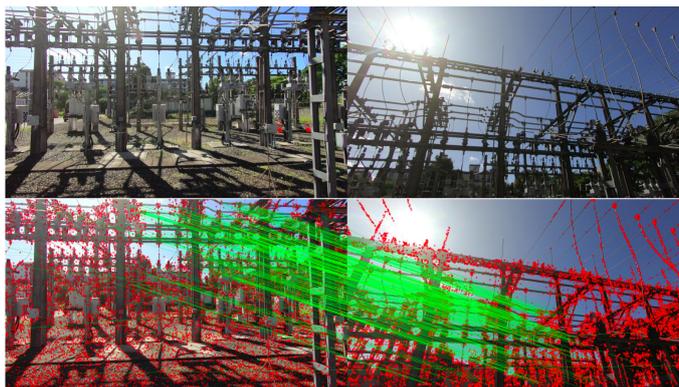


Figura 2: Exemplo de *frames* distintos que observam a mesma região da SE a partir de diferentes pontos de vista (acima) e correspondência dos *keypoints* (abaixo). Fonte: Autor (2022).

Neste ponto já podemos obter a primeira triangulação. Isto é, encontrar a posição tridimensional de cada par de pontos-chaves (cada par de pontos chaves corresponde a um único ponto tridimensional) e também calcular as matrizes de câmera (matriz que leva pontos tridimensionais, pontos do mundo, em pontos da imagem).

Com este processo feito para duas imagens, podemos inserir a próxima imagem. Como o nosso banco de imagens é sequencial (pois foram *frames* extraídos de uma filmagem), poderemos fazer a reconstrução de forma sequencial. Ou seja, sabemos que a próxima imagem tem pontos em comum com a imagem anterior. Em seguida procuramos todos os pontos-chaves desta terceira imagem que estão nas duas anteriores. Isto é, que correspondem a alguns dos pontos-chaves previamente pareados e filtrados. Repetindo o mesmo processo feito anteriormente conseguiremos encontrar a nova matriz de câmera. Este procedimento permitirá olhar para pontos que foram, a priori, descartados, fazendo um melhoramento. Isto é muito importante pois, caso contrário, a cada nova

imagem ficaríamos com menos pontos-chaves e, eventualmente, não teríamos mais pontos chaves em comum.

O processo de reconstrução continua de forma iterativa: uma nova imagem é selecionada dentre os *frames* de entrada, seus pontos-chaves são identificados, e as correspondências destes com os pontos-chaves das imagens anteriores são computados e, por fim, a triangulação de pontos tridimensional é refeita. Cada nova imagem adicionada aumenta a quantidade de pontos tridimensionais conhecidos e melhora numericamente a qualidade dos pontos já existentes, pois adiciona novas medições à cena em reconstrução. Para reduzir os erros acumulados (devido ao processo sequencial) usamos o algoritmo *Bundle Adjustment* [1].

Como resultado final, obtivemos uma nuvem de pontos tridimensionais, que é a reconstrução 3D da SE e, ao mesmo tempo, um mapeamento de cada um destes pontos. Isto é, sabemos como cada um destes pontos é projetado nas imagens (uma vez que temos as matrizes de câmeras). A Figura 3 apresenta uma captura de tela da reconstrução tridimensional da SE realizada. O embasamento matemático envolvido para o processo de uma reconstrução tridimensional pode ser encontrado em [11].

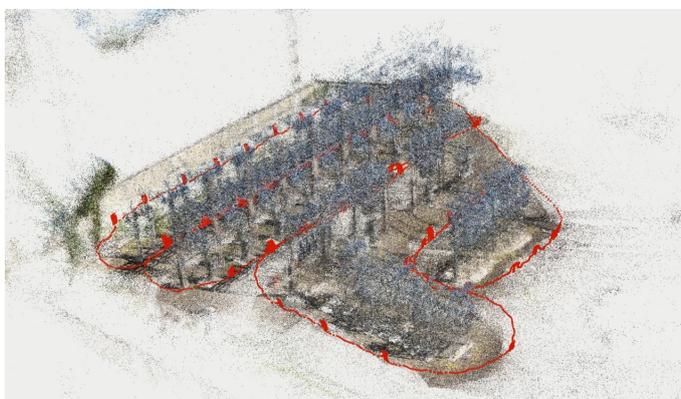


Figura 3: Reconstrução Tridimensional da SE. Fonte: Autor (2022).

Inserção de novas conexões via códigos QR

Para obter uma reconstrução precisa da cena, utilizamos imagens sequenciais provenientes de um vídeo capturado pelo operador da câmera. Ao comparar as imagens para encontrar pontos-chave em comum, comparamos com as vizinhas em uma certa janela, o que é um problema linear, em vez de comparar com todas as outras, o que seria um problema de ordem quadrática. Na Figura 3 podemos ver em vermelho o trajeto realizado pelo operador que movimentou-se pelos corredores, movendo-se lateralmente e verticalmente para garantir a filmagem de todos os equipamentos diversas vezes. Desta forma, o vídeo resultante possui partes que filmam a mesma região da SE, mas em momentos diferentes, o que não é considerado pelo algoritmo sequencial. A fim de tirar proveito destas informações de sobreposição, usamos códigos QR no chão da SE para impor o reconhecimento de conexões não sequenciais, inserindo restrições adicionais ao processo de otimização do *Bundle Adjustment* e diminuindo consideravelmente o erro total da reconstrução.

Amostragem Adaptativa

Para iniciar a reconstrução usamos imagens sequenciais amostradas de um vídeo. Todavia se escolhemos uma amostragem muito grosseira, poderemos acabar perdendo informações importantes

ou criar sequências de imagens sem conexão. Por outro lado, se a amostragem for muito fina, a execução do algoritmo ficará muito pesada. Esta escolha fica ainda mais delicada levando-se em conta que a amostragem adequada não é a mesma em todos os trechos da gravação. Isto é, em locais onde a velocidade angular é grande é necessário uma amostragem mais fina do que em trechos onde o operador da câmera está andando linearmente.

Por isto utilizamos uma amostragem adaptativa. Isto é, primeiro realizamos uma amostragem uniforme e calculamos a matriz essencial conforme processo já descrito. Caso o erro de projeção estiver muito grande (não aceitável), inserimos um novo *frame* da filmagem entre os dois *frames* originais, a fim de melhorar o resultado. Este processo é repetido até que o erro de projeção entre imagens vizinhas seja suficientemente pequeno.

3 Identificação de Equipamentos na Reconstrução

Após realizada a reconstrução 3D da SE, o próximo passo é identificar/localizar na estrutura de dados 3D quais pontos da reconstrução pertencem aos equipamentos da SE. Par tal executamos o seguinte procedimento: identificamos, nos *frames* da reconstrução, aqueles que contêm uma visão clara do equipamento em questão. É importante notar que esses *frames* foram processados pela rede YOLO [9], e que conhecemos todos os equipamentos observados pela reconstrução, bem como suas caixas delimitadoras (*bounding boxes*). Seleccionamos os melhores *frames*, ou seja, aqueles que visualizam os equipamentos mais de perto e que apresentam uma boa variação de ponto de vista. A partir de cada ponto de vista, tomamos a pirâmide formada considerando como vértice o centro da câmera e cujas arestas passam pelos vértices da caixa delimitadora do equipamento em questão. A intersecção de duas ou mais pirâmides resulta num sólido convexo (politopo) que delimita um objeto no espaço tridimensional. Quanto mais pontos de vista, com uma grande variação de posição, melhor será a delimitação do objeto. Note que o sólidos resultantes da intersecção sempre será convexa, e portanto, nunca irá aproximar perfeitamente um objeto não convexo. Porém esta acaba sendo uma ótima aproximação pra envoltória (convexa) do objeto. O politopo resultante da delimitação de cada objeto/equipamento será a representação deste na reconstrução 3D. Atribuímos a cada uma destas envoltórias o nome do equipamento que ela representa e as armazenamos no banco de dados juntamente com a reconstrução 3D da SE.

A Figura 4 ilustra o processo de localização de um objeto na reconstrução 3D. À direita estão os pontos de vista seleccionados. No centro, o volume espacial observado a partir das imagens seleccionadas. Esta interface foi desenvolvida no âmbito deste projeto.

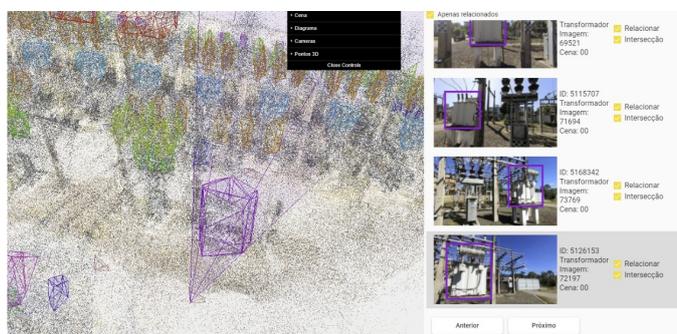


Figura 4: Localização de um objeto na reconstrução tridimensional. Fonte: Autor (2022).

4 Associação Entre Equipamentos Presentes nos Termogramas e os da Reconstrução Tridimensional

As análise das imagens termográficas, sejam estas manuais ou automatizadas irão fornecer um par de imagens (uma ótica e uma térmica) contendo o equipamento possivelmente defeituoso, sendo que a classe do equipamento já foi identificada. Isto é, já sabe-se que o equipamento trata-se de uma chave seccionadora, ou de um transformador, mas não sabe-se exatamente qual. Todavia, para realizar o reparo do equipamento, é necessário identificar o equipamento especificamente. Para isto, iremos encontrar a posição do equipamento na SE através da reconstrução tridimensional.

Como todos os equipamentos da SE já foram previamente localizados na reconstrução tridimensional (através do envólucro detalhado na seção anterior), precisamos realizar a associação entre o equipamento de interesse presente na imagem ótica do termograma e o equipamento correspondente na reconstrução tridimensional.

Para realizar esta associação, o primeiro passo é identificar a posição da câmera (ou, equivalentemente, o ponto de vista) que corresponde a imagem proveniente do termograma. Utilizamos um processo similar ao feito para a reconstrução tridimensional: a partir da imagem (da inspeção), calculamos o SIFT desta para encontrar os pontos chaves. Então procuramos quais das imagens da reconstrução este conjunto de descritores SIFT melhor se encaixa. De posse destas imagem e de seus pontos na reconstrução podemos aplicar um algoritmo de PnP (*Perspective-n-Point*) [4] para encontrar a posição e direção da câmera.

Uma vez conhecido a posição da câmera, verificamos quais equipamentos da SE este ponto de vista é capaz de visualizar. Para tal, projeta-se todos os equipamentos tridimensionais (envoltórias calculadas conforme descrito anteriormente) usando a matriz desta câmera. Se o objeto não estiver na visada desta câmera, ele vai cair em coordenadas fora da imagem. Caso contrário, comparamos sua projeção com o *Bounding Box* do equipamento no imagem. Para fazer isto usamos a OpenGL (*Open Graphics Library*). Esta API (*Application Programming Interface*), além de facilitar a implementação ainda lida com o problema de oclusão de objetos (quando um objeto acaba ocultando outro do ponto de vista da câmera) utilizando o Z-Buffer [2].

Desta forma, conseguimos encontrar o equipamento (o envólucro) na reconstrução tridimensional que corresponde ao equipamento presente na imagem ótica do termograma.

5 Considerações Finais

Neste trabalho foi realizado a reconstrução tridimensional de uma SE a fim de possibilitar a associação automatizada de imagens termográficas de equipamentos elétricos da SE com a posição do respectivo equipamento na SE. Vale mencionar que a SE possui muitos equipamentos do mesmo tipo e que, para o propósito do trabalho, era necessário identificar a posição do equipamento de forma específica.

Embora o processo de reconstrução tridimensional bem como a identificação dos equipamentos nesta reconstrução demande um certo tempo computacional (ordem de dias), ele é realizado apenas uma vez. Por outro lado, a associação dos equipamentos presentes nas imagens com os equipamentos previamente localizados na reconstrução tridimensional é rápido, levando no máximo alguns minutos. Portanto, a abordagem proposta é eficiente no que diz respeito ao tempo demandado para tal, além tanto de realizar a correta localização dos equipamentos a partir dos termogramas.

Apesar das técnicas utilizadas neste trabalho serem conhecidas individualmente, a aplicação em conjunto para o caso de equipamentos de SE é singular. Em particular, o uso das técnicas de inserção de novas conexões utilizando códigos QR e a amostragem adaptativa não são amplamente empregadas no contexto de reconstrução 3D de cenas, o que também diferencia o presente trabalho. É importante ressaltar que esse trabalho foi implementado em uma situação real, demonstrando

sua viabilidade e aplicabilidade prática. Por fim, as ideias aqui empregadas podem ser empregadas em contextos onde haja a necessidade de associar imagens de objetos à sua posição geográfica em um determinado ambiente.

Agradecimentos

Os autores agradecem à UTFPR pelo apoio e infraestrutura disponibilizada para o desenvolvimento desta pesquisa e à COPEL - DIS pelo apoio e financiamento através do projeto PD-2866-0528/2020 - Desenvolvimento de Metodologia para Análise Automática de Imagens Térmicas. Projeto P&D executado pela COPEL-DIS e regulado pela ANEEL.

Referências

- [1] Sameer Agarwal et al. “Bundle adjustment in the large”. Em: **Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5–11, 2010, Proceedings, Part II 11**. Springer. 2010, pp. 29–42.
- [2] Edwin Earl Catmull. “A subdivision algorithm for computer display of curved surfaces”. Tese de doutorado. The University of Utah, 1974.
- [3] Martin A Fischler e Robert C Bolles. “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography”. Em: **Communications of the ACM** 24.6 (1981), pp. 381–395.
- [4] Richard Hartley e Andrew Zisserman. **Multiple view geometry in computer vision**. Cambridge university press, 2003.
- [5] Mohd Shawal Jadin e Soib Taib. “Recent progress in diagnosing the reliability of electrical equipment by using infrared thermography”. Em: **Infrared Physics & Technology** 55.4 (2012), pp. 236–245.
- [6] David G Lowe. “Distinctive image features from scale-invariant keypoints”. Em: **International Journal of Computer Vision** 60.2 (2004), pp. 91–110.
- [7] Pierre Moulon, Pascal Monasse e Renaud Marlet. “Global fusion of relative motions for robust, accurate and scalable structure from motion”. Em: **Proceedings of the IEEE international conference on computer vision**. 2013, pp. 3248–3255. DOI: 10.1109/ICCV.2013.403.
- [8] Trong Phuc Truong et al. “Registration of RGB and thermal point clouds generated by structure from motion”. Em: **Proceedings of the IEEE International Conference on Computer Vision Workshops**. 2017, pp. 419–427.
- [9] Joseph Redmon et al. “You only look once: Unified, real-time object detection”. Em: **Proceedings of the IEEE conference on computer vision and pattern recognition**. 2016, pp. 779–788.
- [10] Johannes L Schonberger e Jan-Michael Frahm. “Structure-from-motion revisited”. Em: **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition**. IEEE. 2016, pp. 4104–4113. URL: <https://colmap.github.io>.
- [11] Eduardo Schumacher. “Geometria projetiva: reconstrução geométrica a partir de múltiplos pontos de vista”. B.S. thesis. Universidade Tecnológica Federal do Paraná, 2021.