

Reconstrução Tridimensional de Ambientes Reais Usando Dados Laser e de Intensidade

Paulo Dias (paulo.dias@ieeta.pt)

Resumo – This article gives an overview of the results obtained during the PhD studies of the author. The work was developed within a cooperation between the IEETA/University of Aveiro and the Joint Research Centre in Ispra, Italy.

The objective was to generate complete, high-resolution three-dimensional models of real world scenes (3D geometry and texture information) from passive intensity images (typically digital photographs) and active range sensors (such as laser range finders).

Depth extraction from intensity images is a complex task. But digital photographs provide additional information about the scenes that can be used to help the 3D modelling process, in particular to define accurate surface boundary conditions. This makes active and passive sensors complementary in many ways and is the fundamental idea that motivates the work presented here.

I. INTRODUÇÃO

A maior parte dos sistemas de reconstrução 3D são baseados ou em dados laser (*range data*) ou em câmaras digitais, sendo ainda pouco habituais sistemas que tentem combinar estes dois sensores[1][2]. A extracção de informação de distância a partir de imagens de intensidade (como fotografias digitais) é complexa. Mas as fotografias fornecem informação adicional sobre os ambientes que pode ser útil durante o processo de modelação em três dimensões. Os sensores activos (laser) e passivos (câmaras) são portanto complementares em vários aspectos (tabela 1), o que motivou o trabalho apresentado neste artigo.

A primeira parte do artigo trata do problema da calibração das imagens, por forma a relacionar os dados oriundos dos sensores de distância activos com as fotografias. Foram desenvolvidas várias ferramentas para tornar este passo mais fácil, independente do utilizador e preciso. No final, com a metodologia proposta, obtêm-se mapas de textura para os modelos baseados em várias fotografias. A geometria do modelo é calculada a partir dos dados laser [3] enquanto que a informação de textura provem de fotografias digitais. Os modelos assim obtidos são como que uma fotografia em três dimensões de alta qualidade de ambientes reais.

Na segunda parte do artigo, apresenta-se uma metodologia mais ambiciosa para a fusão dos dados. As fotografias passam a ser usadas como uma fonte adicional

de informação tridimensional. Informação valiosa para definir com precisão os limites das superfícies (onde a informação dos sensores laser é menos fiável), colmatar lacunas, ou incrementar a densidade de pontos tridimensionais em áreas de interesse.

A figura 1 apresenta um sensor laser Riegl LMS-Z210 [4] que produz uma nuvem de pontos 3D, assim como uma imagem de reflectância e uma fotografia digital do mesmo ambiente adquirida com uma Canon Powershot Pro 70.

	<i>Dados laser</i>	<i>Fotografias</i>
<i>Custo</i>	<i>elevado</i>	<i>reduzido</i>
<i>Aquisição</i>	<i>Complexa, devido ao tamanho do sensor</i>	<i>Fácil, com uma câmara digital</i>
<i>Resolução</i>	<i>Resolução espacial limitada</i>	<i>Fotografias de alta resolução</i>
<i>Mapas de textura</i>	<i>Imagem de reflectância(P&B)</i>	<i>Mapa de textura a cores</i>
<i>Iluminação</i>	<i>Sensor independente da iluminação</i>	<i>Muito dependente da iluminação</i>
<i>Importância da textura</i>	<i>Não é necessária textura</i>	<i>A textura é crucial para obter bons resultados</i>
<i>Processamento 3D</i>	<i>Fornece os dados 3D directamente</i>	<i>Processamento para obter 3D</i>

Tabela 1: Vantagens (sombreado) e limitações dos sensores laser e de intensidade.

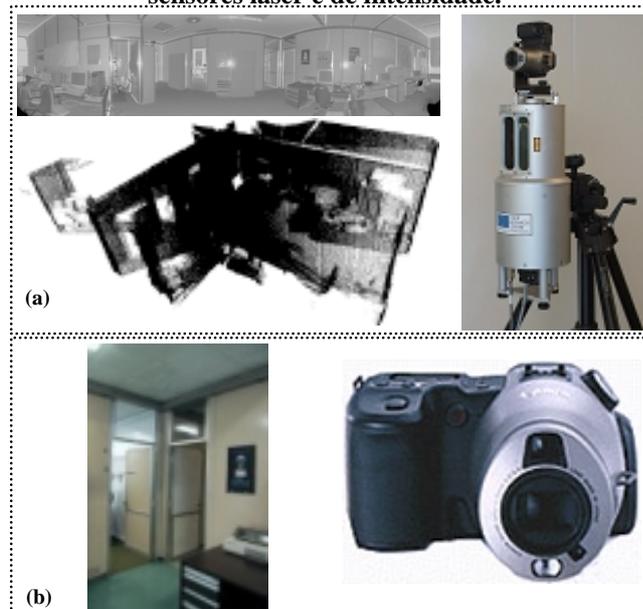


Figura 1: Dados obtidos com o Riegl LMS-Z210 (a) e uma Canon PowerShot Pro 70 (b).

II. CALIBRAÇÃO DA CÂMARA

A. Algoritmo de redimensionamento

Imagens adquiridas com sensores tão diferentes como laser e câmaras fotográficas têm propriedades e resoluções muito distintas. Antes de proceder a qualquer tipo de calibração é necessário realizar um alinhamento inicial por forma a garantir que as características principais das duas imagens (reflectância e fotografia) se sobreponham e que as suas resoluções sejam bastante similares.

Este alinhamento é feito usando o algoritmo apresentado na figura 3. A ideia de base consiste em aproximar a transformação tridimensional entre as duas imagens a uma transformação afim. Os contornos nas imagens (reflectância e intensidade) vão ser usados de uma forma iterativa para avaliar a translação, rotação e escalonamento entre as duas imagens. Esta aproximação só é válida quando as duas imagens são adquiridas a partir de posições bastante próximas umas das outras. Nesse caso, a aproximação afim permite um bom alinhamento inicial entre as duas imagens.

O algoritmo necessita de uma aproximação inicial definida pelo utilizador, o qual deve seleccionar uma área rectangular na imagem de reflectância que cobre aproximadamente a mesma área que a fotografia que se pretende calibrar. O algoritmo implementado detecta e sobrepõe os contornos nas duas imagens. O erro entre os contornos é calculado usando a transformada de distância [5], numa das imagens, para calcular a distância média entre os contornos nas duas imagens. Várias rotações, translações e escalonamentos são então aplicados iterativamente e é seleccionada em cada ciclo a transformação que minimiza a distância entre os contornos. A figura 2 apresenta a sobreposição dos contornos antes (a) e depois (b) de aplicar o algoritmo em duas imagens de um laboratório.

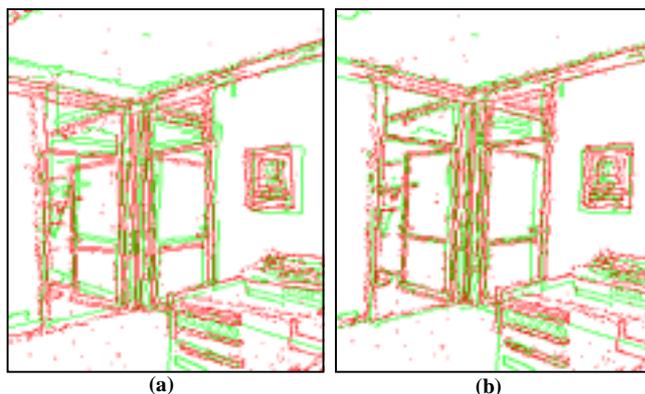


Figura 2: Sobreposição dos contornos antes (a) e depois do algoritmo (b).

A figura 3 apresenta um diagrama de blocos que ilustra o algoritmo de redimensionamento. O resultado é uma sub-imagem de reflectância dimensionada de tal forma que os

contornos entre a fotografia e essa nova imagem sejam aproximadamente sobrepostos.

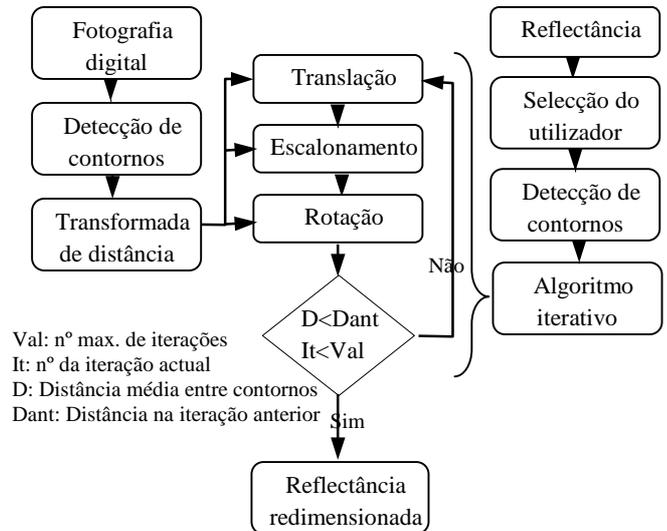


Figura 3: O algoritmo de redimensionamento.

B. Calibração da câmara com o modelo de Tsai

O algoritmo apresentado acima permite obter duas imagens com a mesma resolução e com sobreposição dos principais contornos. Essas imagens vão ser usadas para extrair correspondências entre vários pontos de interesse. Em cada imagem, os pontos de interesse são detectados e comparados usando a correlação cruzada. Obtém-se assim uma lista de pontos correspondentes entre as imagens (ver figura 4). Essas correspondências 2D (na fotografia) e 3D (nos dados laser) vão ser usadas para calcular um modelo da câmara. O modelo usado é o proposto por Tsai [6] e uma técnica RANSAC [7] é usada para aumentar a robustez do sistema. Detalhes sobre a implementação podem ser encontrados em [8].



Figura 4: Correspondências obtidas no exemplo do laboratório.

Uma vez calculado um modelo para a câmara, é possível “re-projectar” a fotografia digital na imagem de reflectância obtendo assim uma nova imagem a cores completamente registada com os dados 3D.

Esse processo pode ser repetido para várias imagens por forma a cobrir toda a área que foi analisada pelo laser, e um processo de *feathering* é usado para permitir transições suaves entre as várias imagens quando estas são reunidas no mapa de textura finalmente obtido.

A figura 5 apresenta os resultados para o laboratório. Em (a) aparece o mapa de textura final obtido através da calibração e re-projectação de 12 fotografias digitais. Em (b) e (c) são apresentadas duas vistas do modelo final com a textura. A figura 6 mostra outro exemplo: o modelo de uma quinta em Laveno, nas margens do “lago Maggiore” no norte da Itália. Neste caso o mapa de textura foi obtido a partir de três fotografias. Os sensores usados nestes dois exemplos são os apresentados na figura 1.

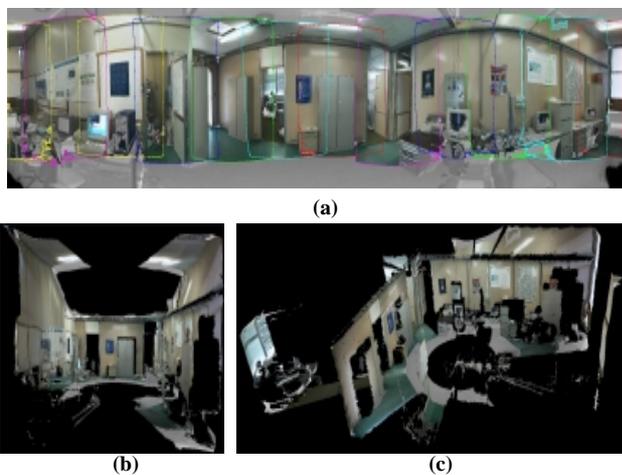


Figura 5: Mapa de textura final obtido a partir de 12 fotografias digitais (a) e duas vistas do modelo (b,c) com a textura obtida com os dados do laboratório.

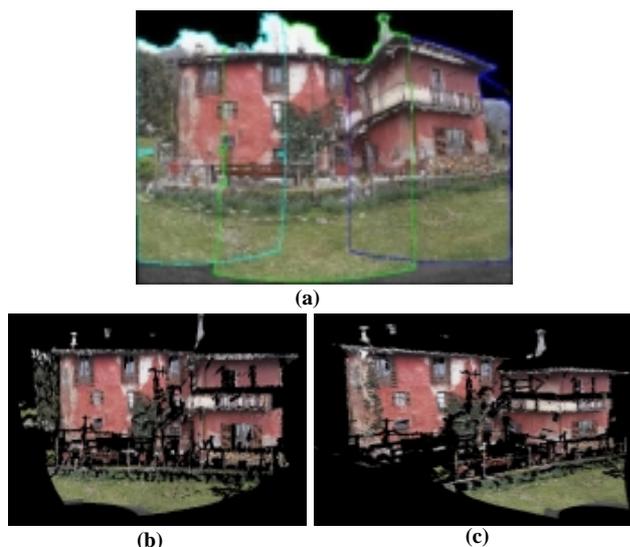


Figura 6: Mapa de textura (a) e duas vistas (b, c) do modelo de uma quinta em Laveno (norte de Itália). Neste caso foram calibradas três fotografias.

III. CORRECÇÃO DE RECTAS 3D BASEADA NUMA SÓ FOTOGRAFIA DIGITAL

No secção anterior, as fotografias digitais adquiridas são usadas apenas como textura para os modelos realizados: toda a informação de geometria é fornecida pelos dados laser.

Os dados laser estão sujeitos a vários tipos de erro. Um erro comum nesse tipo de dados está relacionado com o efeito do ponto misto e ocorre em áreas em que as variações de profundidade são muito elevadas. Esse erro é devido ao facto de o feixe laser não ser um ponto ideal mas uma superfície elíptica cujo diâmetro aumenta com a distância, como ilustrado na figura 7a. Esse fenómeno provoca erros importantes nas discontinuidades e torna irregulares as rectas, como pode ser visto nos contornos da porta do laboratório (figura 7b). Esta secção apresenta uma possível utilização de uma fotografia digital calibrada para minimizar o problema do ponto misto em dados laser e assim melhorar a segmentação de linhas em modelos triangulados.

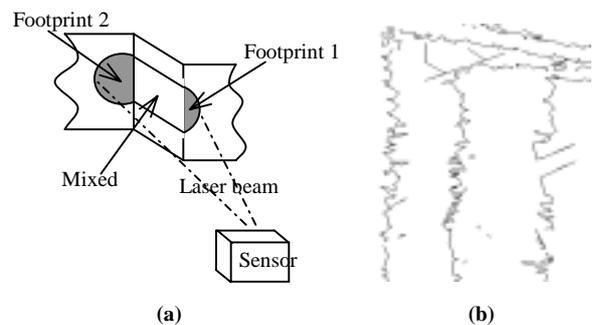


Figura 7: Problema do ponto misto em dados laser e seus efeitos nos contornos de um modelo triangulado.

O processo implementado começa por extrair dos dados de distância os contornos em que existem variações bruscas de profundidade (*jump* e *roof edges*) através do estudo da vizinhança dos pontos em 3D (figura 8b). O algoritmo também detecta os segmentos de recta nas imagens de intensidade que se pretendem usar (figura 8c). A detecção dos segmentos é feita analisando e agrupando os contornos obtidos através de um detector de Canny [9].

O algoritmo vai então detectar, para cada segmento, os contornos 3D que são projectados na sua vizinhança usando o modelo da câmara que se obteve com a técnica apresentada na secção II. A partir dos pontos 3D projectados na vizinhança do segmento 2D, uma possível linha em 3D é definida usando uma técnica RANSAC[7]. Essa linha 3D é usada num segundo passo para detectar os pontos que têm grande probabilidade de pertencer à linha de acordo com as seguintes duas condições:

- A distância entre o ponto e o segmento 3D tem que ser menor que um dado valor.
- Para cada contorno, o ângulo entre a projecção da regressão linear dos pontos 3D e o segmento 2D tem de ser menor que um dado ângulo.

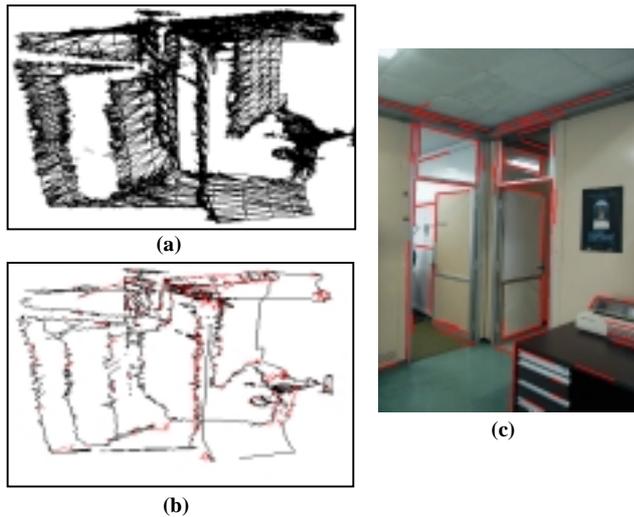


Figura 8: Modelo triangulado de uma parte do laboratório (a). Jump e Roof edges detectados num modelo (b). Fotografia digital calibrada e segmentos detectados.

O passo final do algoritmo consiste, dados os pontos que são considerados como formando um segmento, em calcular o segmento médio em 3D para a qual os pontos vão ser “puxados”. Para esse cálculo foi usado o código apresentado em [10].

A figura 9 apresenta os resultados deste processo no modelo do laboratório. Na figura 9a, são apresentados os contornos iniciais do modelo, enquanto que a figura 9b apresenta os contornos depois da correcção usando a técnica apresentada. É visível a melhoria ao nível da segmentação das rectas, tendo os efeitos do “ponto misto” sido bastante minimizados neste caso.

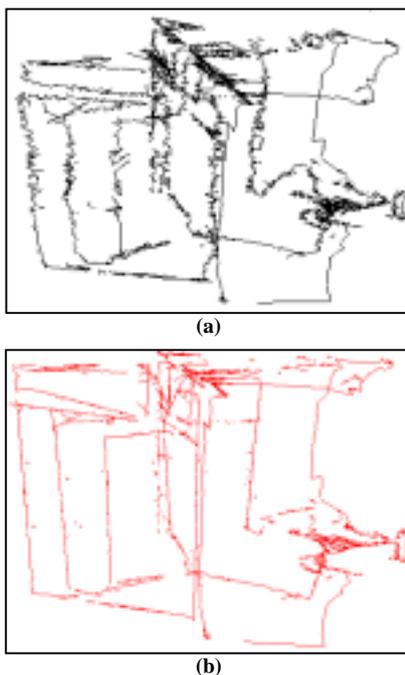


Figura 9: Contornos 3D (edge e roof edges) antes (a) e depois (b) de aplicar o algoritmo para correcção de segmentos baseado numa fotografia digital calibrada.

IV. COMBINAÇÃO DE TRIANGULAÇÃO PASSIVA E DADOS LASER PARA MODELAÇÃO 3D

Na secção anterior, é utilizada uma fotografia digital para melhorar a segmentação de modelos 3D triangulados a partir de dados laser.

No caso de ter disponível várias imagens de uma mesma área, é possível usar técnicas de triangulação para obter directamente informação 3D a partir das fotografias digitais disponíveis.

Nesta secção, pares de fotografias calibradas vão ser usadas para melhorar o modelo da câmara inicial e, depois, introduzir pontos adicionais na nuvem de pontos 3D original obtida a partir dos dados laser. Tal informação pode ser importante para aumentar a resolução dos modelos em áreas de interesse, sem necessidade de adquirir dados laser adicionais, bastando para isso adquirir fotografias das áreas de interesse.

A. Triangulação passiva

A extracção de dados 3D a partir de fotografias não é uma nova área de investigação, e técnicas como o cálculo da matriz fundamental, a rectificação de imagens e a extracção de correspondências entre imagens adquiridas de pontos de vista muito próximos são técnicas já bem conhecidas [11,12,13]. No caso que apresentamos aqui, combinamos estas técnicas com os modelos das câmaras obtidos, permitindo triangular as correspondências a partir dos raios de re-projecção de cada câmara e, dessa forma, obter a localização tridimensional dos pontos a partir das fotografias.



Figura 10: Duas imagens de uma igreja em Laveno, com as respectivas linhas epipolares.

A informação das calibrações iniciais, assim como a correlação cruzada, são usadas para determinar correspondências entre duas imagens da mesma área. A partir desses dados são determinadas a matriz fundamental e as linhas epipolares (figura 10). São então calculadas as imagens rectificadas: duas imagens em que as linhas epipolares correspondem às linhas das imagens. A determinação de correspondências é então simplificada, já

que é possível procurar eventuais correspondências ao longo das linhas das imagens, numa só dimensão.



Figura 11: Imagens rectificadas em que as linhas epipolares aparecem em linhas consecutivas.

A partir das imagens rectificadas (figura 11) é possível calcular de uma forma eficiente correspondências entre as duas imagens usando a correlação cruzada. Dados as correspondências e o modelo das câmaras para as duas imagens, é calculada a intersecção entre os raios de projecção e são trianguladas assim as posições no espaço a partir das correspondências nas duas imagens, como ilustrado na figura 12.

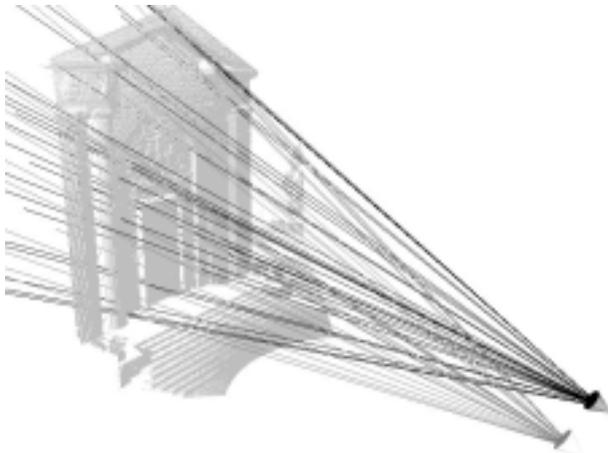


Figura 12: O processo de triangulação passiva.

B. Melhoria da calibração

Os pontos triangulados obtidos pelo método anterior, podem facilmente ser comparados com os pontos medidos inicialmente pelo laser, sendo assim possível ver a qualidade da calibração: quanto maior a distância entre os pontos triangulados e a nuvem de pontos, maior erro existe na calibração.

Essa informação pode ser usada para alinhar os pontos triangulados com a nuvem de pontos, e assim tornar mais precisas as calibrações obtidas inicialmente.

Em cada iteração, os pontos mais próximos dos pontos triangulados são determinados e usados para calibrar novamente as duas imagens. O processo é repetido até que o erro entre os pontos triangulados e a nuvem de pontos fique abaixo de um limiar. As figuras 13 e 14 apresentam este processo. Na figura 13 aparece a nuvem de pontos produzida por um laser Z&F Imager 5003 [14] e os pontos triangulados usando as calibrações iniciais. Na Figura 14, os mesmos dados são apresentados depois de aplicar o nosso algoritmo. É notável a diminuição do erro entre os dados 3D do laser e os pontos triangulados com os novos parâmetros das câmaras.

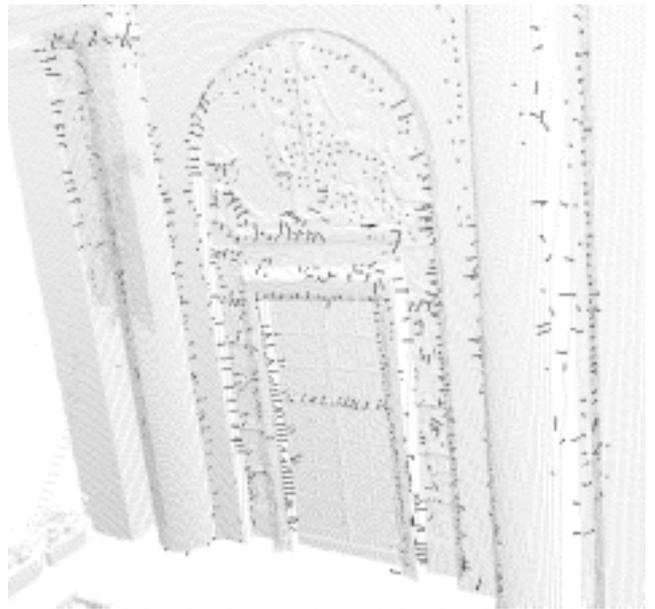


Figura 13: Dados laser (cinzento claro) e pontos triangulados com os parâmetros iniciais das câmaras.

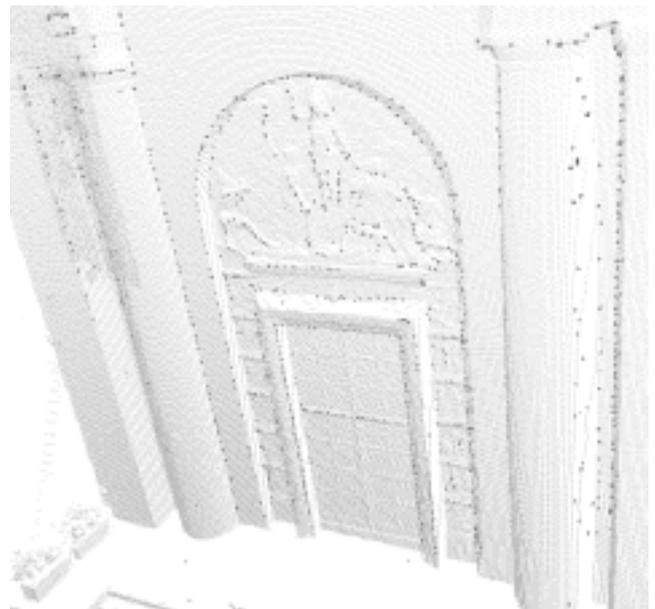


Figura 14: Dados laser (cinzento claro) e pontos triangulados depois do processo de optimização.

C. Adição de pontos 3D nos modelos

As ferramentas desenvolvidas na secção anterior garantem uma boa calibração das fotografias relativamente aos dados laser. Estas ferramentas oferecem também outra possibilidade: permitem adicionar mais pontos tridimensionais à nuvem de pontos original obtida com o laser. Tal possibilidade é particularmente interessante em áreas onde há falta de informação de profundidade (oclusões, falta de dados, etc.) ou em áreas dos modelos particularmente ricas em geometria e onde se deseja mais detalhe. Nestas situações, informação 3D pode ser calculada a partir de pares de fotografias. Esses dados oferecem a vantagem de poder ser adquiridos facilmente, rapidamente e com uma alta resolução.

Neste processo, as imagens rectificadas são usadas para calcular correspondências entre as fotografias. Várias condições são introduzidas para melhorar a pesquisa de correspondências: por exemplo, simetria entre o *matching* nas duas imagens, valores limiares para a correlação cruzada e ordem das correspondências para garantir que esta é a mesma nas duas imagens. Tais condições são já conhecidas e de uso bastante divulgado em algoritmos de *matching* dinâmico [15].

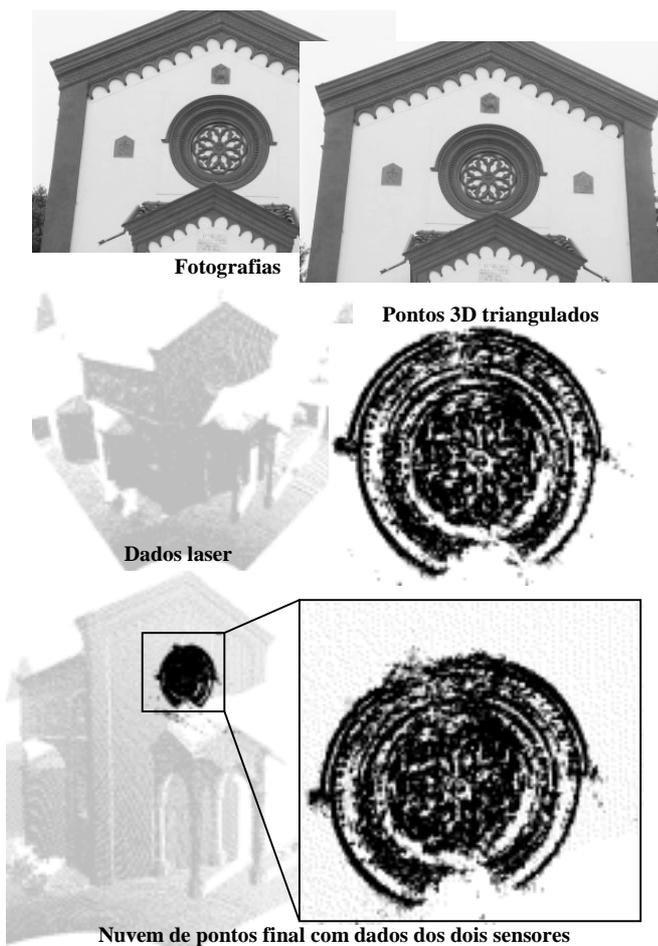


Figura 15: Adição de pontos tridimensionais, a partir de duas fotografias, ao modelo de uma igreja em Barza (Itália).

A figura 15 apresenta um exemplo deste processo com o modelo de uma igreja em Barza, no norte da Itália. Na figura 16, um exemplo similar é apresentado com o modelo da igreja de Laveno.

Nesta situação, são triangulados todos os pontos onde a variação de gradiente é suficientemente elevada para permitir um *matching* com uma certa fiabilidade.

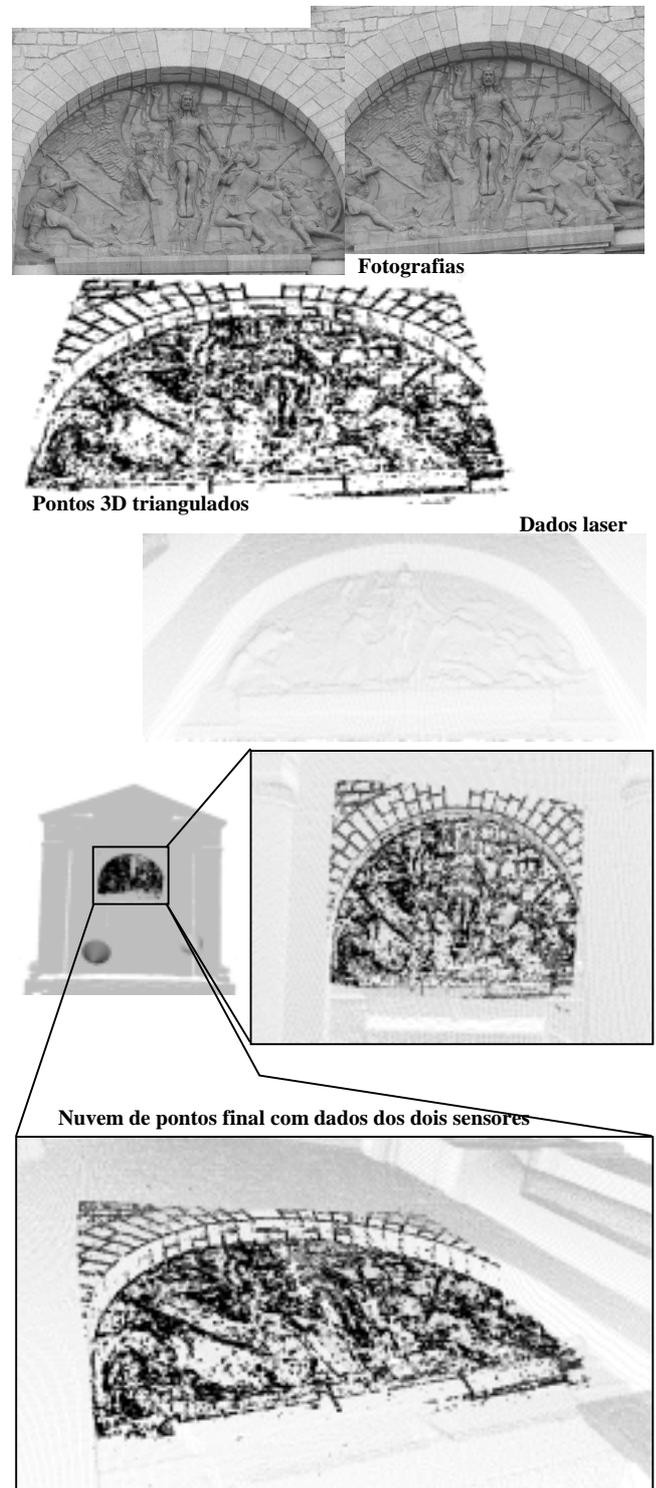


Figura 16: Adição de pontos tridimensionais, a partir de duas fotografias, ao modelo da igreja de Laveno.

V. CONCLUSÃO

Este artigo apresenta alguns dos algoritmos e metodologias desenvolvidos ao longo deste trabalho, para construir modelos tridimensionais de ambientes reais combinando dados laser e de intensidade. Mais detalhes sobre as várias implementações podem ser encontrados em [16].

Os principais resultados deste trabalho foram o desenvolvimento de várias ferramentas para simplificar e melhorar a tarefa de calibração das câmaras com os dados laser, permitindo assim obter mapas de textura para os modelos triangulados com qualidade fotográfica. Num segundo passo foram estudadas varias técnicas para combinar os dados dos dois tipos de sensores. Em particular, fotografias digitais calibradas foram usadas como uma fonte adicional de informação 3D. Este trabalho mostra que os modelos obtidos combinando os dois sensores têm qualidade acrescida, confirmando que a combinação da informação oriunda de vários sensores é uma boa solução para obter modelos tridimensionais mais precisos e de melhor qualidade.

AGRADECIMENTOS

Desejo aqui agradecer ao Professor Francisco Vaz e ao Doutor Vítor Sequeira, orientadores do meu doutoramento, sem os quais não teria sido possível realizar o trabalho aqui apresentado. Os meus agradecimentos também para o Professor João Gonçalves, chefe do sector SIR (Surveillance and Information Retrieval) no Joint Research Centre em Ispra, onde foi desenvolvida a maior parte deste trabalho, e ao Professor A. Sousa Pereira, Director do IEETA, onde foi concluído o meu doutoramento.

Agradeço ainda aos Professores Beatriz Sousa Santos e Joaquim Madeira pelos seus conselhos e comentários sobre este artigo.

Finalmente desejo agradecer à Fundação para a Ciência e Tecnologia, pelo seu apoio financeiro, através da bolsa de investigação PRAXIS XXI/BD/19555/99, que permitiu a concretização deste trabalho.

REFERENCIAS

- [1] S. F. El-Hakim, C. Brenner, G. Roth, "A multi-sensor approach to creating accurate virtual environments", *ISPRS Journal for Photogrammetry and Remote Sensing*, Vol. 53(6), pp. 379-391, December 1998.
- [2] I. Stamos, P. K. Allen, "3-D Model Construction Using Range and Image Data", *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 531-536 (volume I), South Carolina, June 2000.
- [3] V. Sequeira, K. Ng, E. Wolfart, J.G.M. Gonçalves, D. Hogg, "Automated Reconstruction of 3D Models from Real Environments", *ISPRS Journal of Photogrammetry and Remote Sensing (Elsevier)*, vol. 54, pp. 1-22, 1999.
- [4] RiegI GmbH, *Laser Mirror Scanner LMS-Z210- Technical document & user's instruction manual*, November 1999.
- [5] Borgefors G., "A new distance transformation approximating the Euclidean distance". *8th international conference on pattern recognition*. pp. 336-338, 1986.
- [6] R.Y. Tsai, "A versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses", *IEEE Journal of Robotics and automation*, Vol. RA-3, No. 4, pp. 323-344, August 1987.
- [7] M.A. Fischler, R.C. Bolles, "Random Sample Consensus: a paradigm for model fitting with application to image analysis and automated cartography", *Communications of the ACM* 24 (6), pp. 381-395, 1981.
- [8] Reg G. Willson, *Modeling and Calibration of Automated Zoom Lenses*. Ph.D. thesis, Department of Electrical and Computer Engineering, Carnegie Mellon University, January 1994.
- [9] J.F. Canny, "A Computational Approach to Edge Detection". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 6, pp. 679-698, 1986.
- [10] Eberly D.H., *3D Game Engine Design: A Practical Approach to Real-Time Computer Graphics*, Morgan Kaufmann Publishers, September 2000.
- [11] M. Pollefeys, *Self-Calibration and Metric 3D Reconstruction from Uncalibrated Image Sequences*, Ph.D. Thesis, Dept. of Electrical, Katholieke Universiteit Leuven, 1999.
- [12] P.H.S. Torr, D.W. Murray, "The Development and Comparison of Robust Methods for Estimating the Fundamental Matrix", *International Journal on Computer Vision*, Vol. 24, No. 3, pp. 271-300, September 1997.
- [13] Faugeras O.D., *Three-Dimensional Computer Vision: A Geometric Viewpoint*, Cambridge, MA: MIT Press, 1993.
- [14] C. Fröhlich C., M. Mettenleiter, F. Haertl, *Imaging Laser Radar for High-Speed Monitoring of the Environment*, Zoller&Fröhlich technical report 1998.
- [15] I. Cox, S. Hingorani, and S. Rao, "A Maximum Likelihood Stereo Algorithm", *Computer Vision and Image Understanding*, Vol. 63, No.3, pp. 542-567, May 1996.
- [16] P. Dias, *Three dimensional Reconstruction of Real World Scenes Using Laser and Intensity Data*, PhD thesis, University of Aveiro, September 2003.