

Một phương pháp hiệu quả cho quá trình quét bề mặt phản xạ sử dụng Kinect v2

Bùi Văn Biên^{1*}, Bành Tiên Long², Phạm Đức An², Nguyễn Đức Toàn²

¹Bộ môn Cơ khí, Trường Đại học Hải Phòng

²Viện Cơ khí, Trường Đại học Bách khoa Hà Nội

Ngày nhận bài 28/7/2020; ngày chuyển phân biên 31/7/2020; ngày nhận phân biên 28/8/2020; ngày chấp nhận đăng 8/9/2020

Tóm tắt:

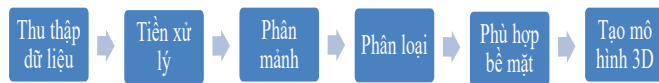
Máy ảnh ToF thường được biết đến là máy ảnh chụp ảnh phạm vi, là một thế hệ cảm biến mới ra đời. Các thiết bị này phù hợp cho việc đo khoảng cách từ thiết bị đến vật thể với tốc độ khung hình cao, được sử dụng ngày càng phổ biến để thu thập dữ liệu 3D cho các ứng dụng robot hay kỹ thuật ngược. Trong bài báo này, các tác giả trình bày một phương pháp hiệu quả để quét các bề mặt phản xạ sử dụng Kinect v2. Phương pháp này dựa trên nguyên tắc của thuật toán ICP. Trong quá trình quét của Kinect v2, đám mây điểm thu được thường không hoàn chỉnh, nó bao gồm các vùng trống dữ liệu. Nghiên cứu này sẽ giải quyết vấn đề vùng trống dữ liệu bằng cách thu thập đám mây điểm theo nhiều hướng khác nhau. Mỗi đám mây điểm sẽ gắn với một hệ trục tọa độ khác nhau. Sau đó, tất cả các đám mây điểm sẽ được căn chỉnh và kết hợp bằng thuật toán ICP. Kết quả thực nghiệm cho thấy hiệu quả của phương pháp đã đề xuất.

Từ khóa: bề mặt phản xạ, đám mây điểm, Kinect v2, quá trình quét 3D, thuật toán ICP.

Chỉ số phân loại: 2.3

Đặt vấn đề

Trong vài thập kỷ gần đây, quá trình số hóa hình dạng của các vật thể (viết tắt là tái tạo ba chiều - 3D) đã được ứng dụng rộng rãi trong nhiều lĩnh vực, từ giải trí tới công nghiệp. Mục tiêu là thu thập dữ liệu 3D của từng điểm trên bề mặt vật thể, tọa độ 3D, độ sâu và véc tơ pháp tuyến. Thông qua một trình tự thực hiện các công việc trong quá trình tái tạo 3D (mô tả trong hình 1), mô hình 3D của vật thể được tái tạo [1].



Hình 1. Trình tự thực hiện tái tạo 3D.

Liên quan đến các ứng dụng trong lĩnh vực chế tạo máy, việc tái tạo 3D nhằm mục đích trích xuất thông tin từ dữ liệu thô được thu thập để tái tạo phiên bản gần nhất của mô hình CAD tham số phù hợp với thiết kế ban đầu của đối tượng. Trong bước “Thu thập dữ liệu” (hình 1), các kết quả khả quan thường được thực hiện với các máy quét 3D thương mại, tuy nhiên giá thành của chúng quá cao, không phù hợp với các doanh nghiệp vừa và nhỏ. Được ra mắt vào năm 2010, Microsoft Kinect (viết ngắn gọn là Kinect v1), thiết bị quét 3D giá rẻ, đã sử dụng công nghệ ánh sáng cấu trúc để thu thập dữ liệu độ sâu. Phiên bản thứ hai của Kinect được phát hành vào năm 2014, giúp cải thiện độ chính xác của

phép đo độ sâu. Kinect v2 áp dụng kỹ thuật thời gian truyền sóng (ToF) để thu thập dữ liệu độ sâu. Kỹ thuật này cũng giúp đạt được thông tin độ sâu với độ phân giải tốt hơn trong khi hạn chế nhiễu từ các nguồn bên ngoài [2]. Đây chính là tiềm năng lớn trong việc sử dụng Kinect v2 cho việc thu thập dữ liệu trong quá trình tái tạo bề mặt 3D.

Kể từ khi xuất hiện, Kinect đã là một đối tượng cho nhiều phương pháp đã được phát triển với hy vọng đưa ra đánh giá kỹ lưỡng về độ chính xác của nó. Một mô hình toán cho dữ liệu độ sâu thu thập từ Kinect đã được đề xuất [3]. Trong đó, một nghiên cứu sâu về ảnh hưởng của các tham số đến độ chính xác của dữ liệu độ sâu thông qua phân tích sai số lý thuyết đã được giới thiệu. Để tính toán dữ liệu 3D, các tham số nội và ngoại của quá trình hiệu chuẩn, chẳng hạn như tiêu cự, điểm góc, hệ số méo ống kính, chiều dài cơ sở và khoảng cách mẫu tham chiếu được sử dụng trong mô hình đã được đề xuất. Kết quả thí nghiệm đã chứng minh rằng, sai số ngẫu nhiên của dữ liệu đo độ sâu tỷ lệ thuận với sự gia tăng khoảng cách giữa máy ảnh và đối tượng đo.

Một đánh giá toàn diện về Kinect v2 cho mục đích tái tạo 3D đã được thực hiện bởi Lachat và các cộng sự [4]. Một số thí nghiệm đã góp phần làm nổi bật các sai số phát sinh từ môi trường và các thuộc tính của vật thể được chụp, cũng như các sai số liên quan đến bản thân thiết bị. Qua những đánh giá nguồn gây sai số, hiệu chuẩn hình học cũng như

*Tác giả liên hệ: Email: bienbv80@dhhp.edu.vn

An efficient method for reflective surface scanning using Kinect v2

Van Bien Bui^{1*}, Tien Long Banh²,
Duc An Pham², Duc Toan Nguyen²

¹Department of Mechanical Engineering, Hai Phong University

²School of Mechanical Engineering, Hanoi University of Science and Technology

Received 28 July 2020; accepted 8 September 2020

Abstract:

ToF cameras, also identified as range imaging cameras, are a recent generation of sensors. As they are desirable for measuring distances to objects at a high frame rate, such sensors are increasingly used for 3D acquisitions, and more generally for applications in robotics or reverse engineering. In this study, the authors presented an efficient method for reflective surface scanning using Kinect v2. The approach is based on the principles of the ICP (Iterative Closest Point) algorithm. During the Kinect v2 scanning, the obtained point cloud is incomplete and includes flying pixels. The authors solve the flying pixel problem by capturing point clouds as multiple viewpoints. Each point cloud belongs to each coordinate. Then, all of the point clouds must be aligned and combined using the ICP algorithm. The experimental results show that the quality of the method is significantly improved.

Keywords: ICP algorithm, Kinect v2, point cloud, reflective surface, 3D scanning.

Classification number: 2.3

độ sâu đã được thực hiện để làm giảm ảnh hưởng của sai số đến độ chính xác của Kinect. Cuối cùng, bề mặt 3D của một mảnh khảo cổ đã được xây dựng, qua đó chứng minh được khả năng ứng dụng và độ chính xác của máy ảnh Kinect trong việc tái tạo mô hình 3D của các vật thể.

Đối với các hệ thống sử dụng nhiều Kinect v2, một phương pháp mới đã được phát triển bởi Yang và cộng sự [5] để cải thiện độ chính xác trong đo độ sâu. Công trình này tập trung vào việc đánh giá độ chính xác độ sâu của Kinect v2. So với Kinect 1, hiệu năng trên phần cứng theo các tham số kỹ thuật đã ban hành được trình bày kỹ lưỡng. Nhiều thuộc tính quan trọng của máy ảnh Kinect v2 sử dụng trong thực tế được nghiên cứu, như phân bố tính chính xác, độ phân giải độ sâu, entropy độ sâu, nhiễu tại giao tuyến của các bề mặt cũng như nhiễu cấu trúc. Các kết quả góp phần thực hiện các thí nghiệm kiểm tra và đưa ra độ chính xác tốt cho Kinect v2 sau khi xác định vị trí đối tượng trong các khu vực phù hợp.

Cũng trong nghiên cứu trước đây của nhóm tác giả [1], ảnh hưởng của tính phản xạ trong quá trình thu thập dữ liệu 3D đã được trình bày. Ba mẫu vật liệu đã được nghiên cứu là nhôm, thép và nhựa. Qua thí nghiệm cho thấy, ảnh hưởng của tính chất bề mặt đối tượng đến quá trình quét sử dụng Kinect v2 là rất rõ ràng. Chính vì vậy, trong nghiên cứu này, phương pháp nhằm giảm tác động của phản xạ bề mặt đối tượng được quét đối với sự hình thành đám mây điểm trong quá trình tái tạo bề mặt 3D của các chi tiết cơ khí sẽ được đề xuất và đánh giá.

Nội dung nghiên cứu

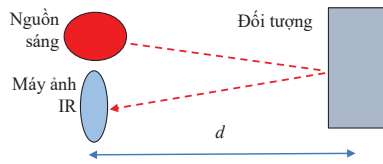
Thu thập dữ liệu đám mây điểm

Kinect v2 bao gồm hai máy ảnh, là máy ảnh hồng ngoại (IR) và máy ảnh RGB. Trong thiết bị này cũng có ba nguồn ánh sáng IR, mỗi nguồn sáng tạo ra một sóng được điều chế với biên độ khác nhau. Thông số kỹ thuật của thiết bị được minh họa trong bảng 1.

Bảng 1. Thông số kỹ thuật của Kinect v2.

Độ phân giải máy ảnh RGB	1920x1080 pixels
Độ phân giải máy ảnh IR	512x424 pixels
Tỷ lệ khung hình	Đến 30 Hz
Thị trường máy ảnh IR	70x60 độ
Thị trường máy ảnh màu	84x54 độ
Phạm vi làm việc	Từ 0,5 đến 4,5 m

Để xác định giá trị độ sâu, Kinect v2 sử dụng công nghệ ToF quang. Dựa trên kết quả đo thời gian để sóng ánh sáng truyền từ nguồn sáng hồng ngoại đến vật thể và sau đó quay trở lại máy ảnh hồng ngoại, giá trị độ sâu sẽ được xác định (hình 2).



Hình 2. Nguyên lý làm việc ToF.

Gọi d là khoảng cách từ Kinect v2 đến vật thể, dựa trên ánh sáng đã điều chế, sử dụng hệ thống ToF gián tiếp thay vì đo trực tiếp thời gian di chuyển, độ lệch pha $\Delta\varphi$ giữa tín hiệu phát và tín hiệu thu sẽ được đo. Khoảng cách ước tính d được tính theo công thức sau:

$$d = \frac{\Delta\varphi}{4\pi f} \cdot c \quad (1)$$

trong đó: f là tần số điều chế, c là tốc độ ánh sáng trong không khí.

Mỗi máy ảnh Kinect v2 đều có các tham số nội riêng phụ thuộc vào cảm biến và ống kính. Các tham số này đã được hiệu chỉnh tại nhà máy và lưu trữ ở bộ nhớ trong của thiết bị. Sử dụng công cụ Kin2 [6] phát triển trên Matlab, các giá trị của tham số nội được sử dụng trong bài viết này được thu thập và hiển thị trong bảng 2.

Bảng 2. Các tham số nội của máy ảnh IR.

Tham số nội		Giá trị (Pixels)
Độ dài tiêu cự	F_x	362,8785
	F_y	362,8785
Điểm góc	C_x	253,8325
	C_y	204,4191
Độ méo hướng kính	K_1 (2 nd)	204,4191
	K_2 (4 th)	-0,2698
	K_3 (6 th)	0,0888

Các bản đồ độ sâu của đối tượng sau đó được chuyển đổi thành các đám mây điểm 3D sử dụng các tham số nội của máy ảnh hồng ngoại, dữ liệu độ sâu thu được và mô hình máy ảnh theo quan hệ hình chiếu phối cảnh. Mỗi điểm ảnh $p(u,v)$ trong bản đồ độ sâu được chuyển đổi thành một điểm $P(X,Y,Z)$ trong không gian thực và biểu diễn bởi đám mây điểm 3D, trong đó gốc tọa độ của đám mây điểm được tạo ra đặt tại vị trí máy ảnh IR của Kinect v2. Các tọa độ X và Y của điểm P tương ứng với từng điểm ảnh p trong bản đồ độ sâu và được tính bằng các công thức:

$$X = \frac{u - C_x}{F_x} \cdot Z \quad (2)$$

$$Y = \frac{v - C_y}{F_y} \cdot Z \quad (3)$$

trong đó: Z là giá trị cường độ của điểm ảnh $p(u,v)$ trong bản đồ độ sâu.

Ghép đám mây điểm bù vùng phản xạ

Qua thực nghiệm và công bố của nhóm tác giả [1], đám mây điểm thu được của các vật liệu có hệ số phản xạ bề mặt lớn như nhôm sẽ có những vùng không có dữ liệu (flying fixels). Có thể nhận thấy rằng, phản xạ bề mặt phụ thuộc vào góc chụp của đối tượng. Các vùng không có dữ liệu trên bề mặt thu được có liên quan đến góc giữa tia sáng IR chiếu từ nguồn sáng và góc thu của máy ảnh IR. Máy ảnh Kinect v2 chụp được các vùng khác nhau của đối tượng, dựng được tọa độ thành phần, sau đó tổng hợp thành các tọa độ toàn thể. Các đám mây điểm của đối tượng sẽ được ghép, vùng bị phản xạ của đám mây điểm này sẽ được bù với vùng không bị phản xạ ở đám mây điểm khác. Để ghép các đám mây thành phần thành một đám mây của đối tượng hoàn chỉnh, việc cần làm là phải lựa chọn sao cho vùng bị phản xạ ở đám mây điểm này trùng với một vùng không chịu ảnh hưởng của phản xạ bề mặt ở một đám mây điểm khác.

Khó khăn lớn nhất trong việc ghép các đám mây dữ liệu thu được chính là mỗi lần quét được thực hiện trên một hệ tọa độ khác nhau. Trong một số trường hợp, đối tượng được di chuyển đến các vị trí khác nhau trong suốt quá trình quét và được kiểm soát bởi máy tính. Khi đó yêu cầu kết quả thu được tại các vị trí quét khác nhau phải có cùng độ chính xác. Đây là trường hợp khi đối tượng đặt trên một bàn xoay hoặc dịch chuyển tịnh tiến và được điều khiển bằng máy tính. Trong bài báo này, thuật toán ICP [7] được sử dụng để ghép hai đám mây điểm. Giả sử hai đám mây điểm lần lượt được gọi là D và M , đồng thời số điểm thuộc D và M tương ứng là N_D và N_M . Công thức của phép chuyển đổi hệ tọa độ như sau [7]:

$$D = R \cdot M + T \quad (4)$$

trong đó R và T lần lượt là ma trận quay và ma trận tịnh tiến cần tìm để hai đám mây điểm có thể ghép với nhau với sai số được tính theo phương pháp bình phương cực tiểu sau:

$$E(R, T) = \sum_{i=1}^{N_M} \|(R \cdot M_i + T) - D_j\|^2 \quad (5)$$

trong đó (M_i, D_j) là cặp điểm tương đồng với nhau trong hai đám mây điểm. Giữ nguyên giá trị M_i , điểm tương đồng D_j được tính theo công thức dưới đây:

$$M_v = \arg \min_{k \in \{1, 2, \dots, N_D\}} \|(R \cdot M_i + T) - D_k\|^2 \quad (6)$$

Quy trình thực hiện ghép đám mây điểm bù vùng trống dữ liệu do phản xạ bề mặt sử dụng thuật toán ICP bao gồm các bước: chọn các điểm tương đồng trong đám mây điểm

tham chiếu; liên kết các điểm này bởi điểm lân cận gần nhất với những điểm tương đồng trong đám mây điểm tiếp theo; ước lượng các ma trận quay và chuyển vị trong phép chuyển đổi sử dụng thuật toán ICP; sử dụng các ma trận ước lượng R và T đã xác định ở trên để chuyển đổi điểm giữa hai đám mây điểm; lặp lại các bước trên cho đến khi các điều kiện dừng của thuật toán ICP được đáp ứng.

Kết quả và thảo luận

Mô tả thí nghiệm

Mục đích của thí nghiệm là nhằm khảo sát khả năng của phương pháp ghép đám mây điểm bù vùng trống dữ liệu trong việc giảm ảnh hưởng của phản xạ bề mặt. Kết quả của thí nghiệm sẽ được đánh giá thông qua hai yếu tố, thứ nhất là hình dạng ban đầu của vật mẫu có được biểu diễn rõ ràng qua các điểm thu được hay không và thứ hai là mật độ điểm trong dữ liệu đám mây điểm bị ảnh hưởng như thế nào.

Để thực hiện khảo sát ghép đám mây điểm bù vùng trống dữ liệu, một chi tiết vật liệu nhôm có kích thước 120x120x14 mm được gia công trên máy phay CNC Manford MCB-850. Phần mềm KSCAN3D được sử dụng để thu nhận dữ liệu các đám mây điểm. Sau đó, việc ghép đám mây điểm theo thuật toán đề xuất ở trên sẽ được triển khai bằng chương trình máy tính bằng mã Matlab.

Thí nghiệm được tiến hành lần lượt với 4 vị trí theo dịch chuyển tịnh tiến, minh họa trong hình 3 và 36 vị trí theo dịch chuyển quay, minh họa trong hình 4. Trong quá trình chuyển động, các vật mẫu được đặt trên hệ thống cơ khí hỗ trợ nhằm đảm bảo sai lệch vị trí là thấp nhất có thể sau mỗi lần vật mẫu dịch chuyển.



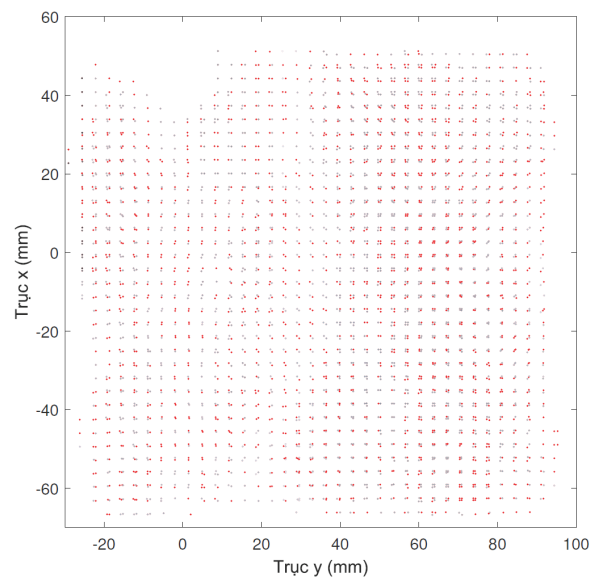
Hình 3. Hệ thống hỗ trợ quá trình thực nghiệm theo dịch chuyển tịnh tiến.



Hình 4. Hệ thống cơ khí hỗ trợ quá trình thu thập dữ liệu theo dịch chuyển quay.

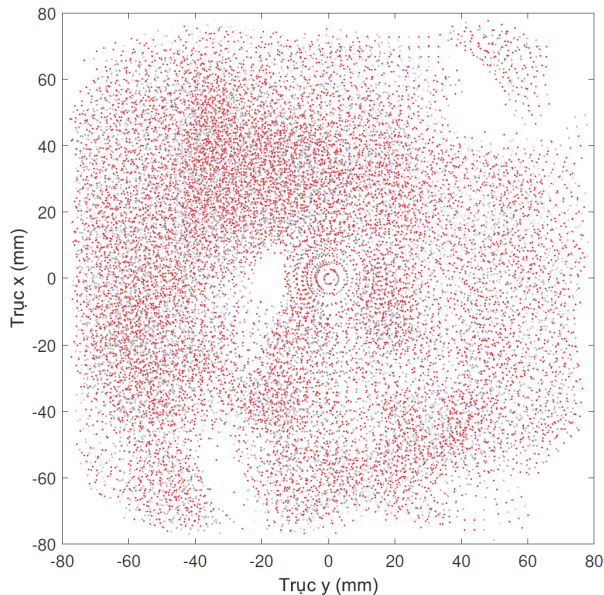
Kết quả và đánh giá

Sau khi thu nhận được toàn bộ 4 đám mây điểm ở 4 vị trí khảo sát theo dịch chuyển tịnh tiến, trước khi ghép, các đám mây điểm được lọc bởi khối lọc 130x130x10 mm để loại bỏ các điểm ở mặt phẳng tham chiếu và môi trường xung quanh. Kết quả ghép đám mây điểm bù vùng trống dữ liệu khi vật mẫu dịch chuyển tịnh tiến được minh họa trong hình 5. Đám mây điểm này thể hiện khá rõ hình dạng ban đầu của vật mẫu. Qua so sánh với các đám mây điểm ở 4 vị trí khảo sát trước khi ghép, các vùng trống dữ liệu ở khu vực trung tâm của bề mặt quét tại các đám mây điểm khác nhau đã được bù lẫn nhau.



Hình 5. Đám mây điểm sau khi ghép để bù vùng trống dữ liệu theo dịch chuyển tịnh tiến.

Đối với dịch chuyển quay, sau khi thu nhận được toàn bộ 36 đám mây điểm, các đám mây điểm này cũng được lọc tương tự như trong dịch chuyển tịnh tiến. Kết quả ghép đám mây điểm bù vùng trống dữ liệu theo dịch chuyển quay được biểu diễn ở hình 6. Đám mây điểm này còn một số vùng trống dữ liệu ở khu vực trung tâm và chưa thể hiện rõ hình dạng ban đầu của vật mẫu. Nguyên nhân là do mỗi vị trí khảo sát cách nhau 10° nên quá trình số hóa không thu thập đủ dữ liệu cần thiết.



Hình 6. Đám mây điểm sau khi ghép để bù vùng trống dữ liệu theo dịch chuyển quay.

Như vậy, việc đăng ký đám mây điểm để bù vùng trống dữ liệu sẽ phù hợp với quá trình số hóa bề mặt khi các vùng trống dữ liệu không quá lớn, các vùng trống không ở biên giới của bề mặt. Hơn nữa, phương pháp ghép đám mây điểm thực hiện tính toán trong thời gian khoảng 5 phút cho quá trình ghép đám mây điểm. Nếu chi tiết có kích thước lớn, mật độ điểm tăng lên, thời gian tính toán sẽ tăng rất nhanh.

Vì vậy, có thể áp dụng các thuật toán xác định biên dạng của vùng trống dữ liệu để giảm thời gian tính toán trong quá trình ghép đám mây điểm.

Kết luận

Qua kết quả thực nghiệm cho thấy, phương pháp ghép đám mây điểm bù vùng trống dữ liệu do bề mặt phản xạ của sản phẩm cơ khí sau gia công cắt gọt gây ra trong quá trình quét 3D bằng thiết bị Kinect v2 cho kết quả khả quan. Các đám mây điểm sau khi ghép đều thể hiện rõ hình dạng ban đầu của vật mẫu. Đồng thời mật độ đám mây điểm sau khi ghép cũng tăng lên đáng kể. Nhằm đánh giá hiệu quả của phương pháp đã đề xuất ở trên, trong nghiên cứu tiếp theo, các tác giả sẽ thực hiện so sánh với một số phương pháp giảm ảnh hưởng của phản xạ bề mặt khác.

TÀI LIỆU THAM KHẢO

- [1] V.-B. Bui, T.-L. Banh, and D.-T. Nguyen (2019), "Improving the depth accuracy and assessment of Microsoft Kinect v2 towards a usage for mechanical part modeling", *J. Korean Soc. Precis. Eng.*, **36**, pp.691-697.
- [2] H. Sarbolandi, D. Lefloch, and A. Kolb (2015), "Kinect range sensing: Structured-light versus time-of-flight kinect", *Comput. Vis. Image Underst.*, **139**, pp.1-20.
- [3] K. Khoshelham and S.O. Elberink (2012), "Accuracy and resolution of Kinect Depth data for indoor mapping applications", *Sensors*, **12**, pp.1437-1454.
- [4] E. Lachat, H. Macher, T. Landes, and P. Grussenmeyer (2015), "Assessment and calibration of a RGB-D camera (Kinect v2 Sensor) towards a potential use for close-range 3D modeling", *Remote Sens.*, **7**, pp.13070-13097.
- [5] L. Yang, L. Zhang, H. Dong, A. Alelaiwi, and A. El Saddik (2015), "Evaluating and improving the depth accuracy of Kinect for Windows v2", *IEEE Sens. J.*, **15**, pp.4275-4285.
- [6] J.R. Terven and D. M. Córdova-Esparza (2016), "Kin2: A Kinect 2 toolbox for MATLAB", *Sci. Comput. Program.*, **130**, pp.97-106.
- [7] K. Low (2004), *Linear least-squares optimization for point-to-plane ICP surface registration*, Univ. North Carolina.