A Dynamic 3D Foot Reconstruction System

Ali K. Thabet, Emanuele Trucco, Joaquim Salvi, Weijie Wang, and Rami J. Abboud

Abstract— Foot problems are varied and range from simple disorders through to complex diseases and joint deformities. Wherever possible, the use of insoles, or orthoses, is preferred over surgery. Current insole design techniques are based on static measurements of the foot, despite the fact that orthoses are prevalently used in dynamic conditions while walking or running. This paper presents the design and implementation of a structured-light prototype system providing dense three dimensional (3D) measurements of the foot in motion, and its use to show that foot measurements in dynamic conditions differ significantly from their static counterparts. The input to the system is a video sequence of a foot during a single step; the output is a 3D reconstruction of the plantar surface of the foot for each frame of the input. Engineering and clinical tests were carried out for the validation of the system. The accuracy of the system was found to be 0.34 mm with planar test objects. In tests with real feet, the system proved repeatable, with reconstruction differences between trials one week apart averaging 2.44 mm (static case) and 2.81 mm (dynamic case). Furthermore, a study was performed to compare the effective length of the foot between static and dynamic reconstructions using the 4D system. Results showed an average *increase of 9 mm* for the dynamic case. This increase is substantial for orthotics design, cannot be captured by a static system, and its subject-specific measurement is crucial for the design of effective foot orthoses.

I. Introduction

This paper presents a novel system to reconstruct the 3D shape of the human foot in motion (4D data) to support a more accurate design of insoles for foot treatment. We use the system to show that foot measurements in dynamic conditions differ sufficiently from their static counterparts to have significant consequences for the design of effective foot orthoses. This is the motivation for our work.

The foot provides a crucial contribution to the balance and stability of the musculoskeletal system [1]. Foot problems are common among a wide group of people, ranging from simple disorders through to complex diseases and joint deformities. When planning treatment, a conservative trend is considered first, with surgery being a last resort. Conservative treatments consist normally in the design of insoles, or *orthoses*, which are intended to ease pain, improve mobility, and correct deformities. Insoles can be designed using conventional casting methods or more advanced Computer Aided Design

Ali K. Thabet is with the Institute of Motion Analysis and Research, University of Dundee, UK akthabet@dundee.ac.uk

Emanuele Trucco is with the School of Computing, University of Dundee, UK manueltrucco@computing.dundee.ac.uk

Joaquim Salvi is with the Department of Computer Architecture and Technology, University of Girona, Spain qsalvi@silver.udg.edu

Weijie Wang is with the Institute of Motion Analysis and Research, University of Dundee, UK wwang@dundee.ac.uk

Rami J. Abboud is with the Institute of Motion Analysis and Research, University of Dundee, UK rjabboud@dundee.ac.uk

(CAD) and surface reconstruction systems. Currently, both design solutions are based on static measurements, despite the fact that orthoses are primarily design to be used during dynamic activities.

Three-dimensional surface reconstruction techniques have been used to reconstruct the shape of the foot while the patient is in a static position. Unfortunately static measurements appear increasingly inadequate; several studies show that treatments based on insoles designed around static foot models are far from optimal [2], [3], [4].

We designed and implemented a *4D Foot Reconstruction System* (henceforth 4DFRS) achieving 3D measurements of the foot in motion via coded structured light. A light source projects a specially designed pattern of coloured stripes upon the surface of the foot. A video camera then acquires frames of the illuminated foot taking a step. For each frame, the camera position of the coloured stripes is used to reconstruct the 3D foot surface by calibrated triangulation. The result is a 4D model of the foot, i.e., a series of surfaces that describe capturing the shape of the foot surface during motion.

II. Related Work

The incorporation of digitized 3D measurements in the field of orthotic design provides the foot specialist with an alternative to plaster casts. Using computerized 3D foot models increases production efficiency substantially [5]. Several commercial static foot scanners are currently available [6], [7], with prices ranging from 6,000 to over 20,000 USD.

In an effort to provide cost efficient solutions to acquire the 3D shape of the foot, several researches have designed foot scanners that implement off-the-shelf hardware [8], [9], [10]. These techniques work well in static conditions, but none of them provides a dynamic reconstruction of the foot. Recent work by Wang *et al.* presents a multi-camera dynamic foot reconstruction system [11]. The authors use *Principal Component Analysis (PCA)* to fit their reconstructed data into a generic foot model obtained from a foot shape database. This system cannot reconstruct the plantar surface of the foot while our more clinically oriented application provides dynamic plantar surface reconstructions (which are essential in designing insoles), in addition to experimentally showing the significant difference between static and dynamic measurements.

Codert *et al.* developed a technique to reconstruct the surface of the whole foot in motion [12]. The authors implemented a stereo vision technique using three pairs of stereoscopic camera. Since the foot images contain insufficient information to solve the correspondence problem [13], the authors considered two options; one was to cover the

foot with a sock while the second involved spraying the foot surface with paint in order to embed on it a random grey level pattern. These methods yielded good outcomes, but invasive factors like socks or paint are undesirable in a clinical environment.

III. METHODOLOGY

A. System overview

3D reconstruction of the foot with passive stereo is limited by the lack of texture on the foot's surface. An effective solution creates artificial texture by projecting light onto the target surface [14], [15], [16]. The light pattern projected varies across techniques; for instance, *Coded Structured Light (CSL)* projects a coded sequence of coloured stripes to make stripe location unambiguous [16]. 4DFRS adopts this technique within a camera-projector set-up.

The architecture of 4DFRS has three main modules: *Calibration*, *Sequence acquisition*, and *3D reconstruction*.

B. Calibration

The equipment was calibrated by modelling both camera and projector using the well-known *pinhole model* [17] as a projector can be modelled as an inverse camera [13].

Calibration requires the acquisition of corresponding 3D and 2D point pairs. This is achieved with the use of a specially designed calibration pattern. Corresponding pairs are fed to a calibration algorithm in order to obtain the parameters of the camera and projector. We adopted the Faugeras-Toscani calibration algorithm in 4DFRS [18].

C. Sequence acquisition

A comprehensive review of CSL techniques by Salvi *et al.* [16] shows that techniques using spatial neighbourhood coding are the most suitable for dynamic surface reconstruction [19], [14]. The 4DFRS uses the simple yet accurate approach by Pages *et al.* [13], i.e., spatial-coded/peak-based structured light. Fig. 1 illustrates the pattern used in the foot reconstruction system. It consists of 64 coloured stripes with black bands between each pair of consecutive stripes. The arrangement of stripes is based on a De Bruijn sequence [16] of four colours and window property of three, meaning that any three consecutive stripes form a unique colour sequence within the pattern. This enables robust correspondence estimation. The pattern illuminates the foot surface taking a step as the video camera records a sequence of frames. These frames are then processed sequentially to reconstruct the shape of the foot surface.

D. 3D reconstruction

Correspondences between the image stripes and the original pattern are calculated. This process occurs in two stages. Stage one consists of locating the centres of the coloured stripes on the captured image; stage two compares the coded patterns surrounding the centres in order to determine correspondences. Details on the initial stage can be found in [13].
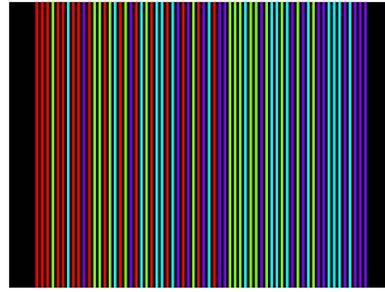

Fig. 1. CSL pattern used in the 4DFRS. The pattern shows a De Bruijn sequence of four colours (red, green, cyan, and magenta), with every 3 consecutive colours being unique within the sequence

Following Zhang *et al.* [14], the correspondence between image stripes and pattern stripes was solved using single-pass dynamic programming, a well-established approach in solving the correspondence problem in structured-light systems and passive stereo [13].

The 3D co-ordinates of surface points are calculated by triangulation. The back-projection ray through the centre of the camera reference frame and a stripe pixel is intersected with the plane defined by the centre of the projector reference frame and the pattern stripe corresponding to the pixel.

IV. IMPLEMENTATION AND TESTING

A. System implementation and experiment design

The 4DFRS prototype was set up in the gait laboratory of the Institute of Motion Analysis and Research (IMAR) of the University of Dundee. To minimize colour distortions in the projected and observed stripe patterns, we selected a three-chip camera (Panasonic HDC-H200 high-definition, with 1080i HD resolution) and 3-LCD projector (Sony VPL-EX4). The prototype was situated in a recess below a dedicated walkway, covered by a transparent slab of toughened glass as a walking platform. The stand-off distance (system to target) was approximately 100 cm. The depth (range) of the reconstruction workspace is approximately 10 cm. We verified experimentally that the transparent platform did not alter the shape, position or colour of the projected stripe patterns, i.e., the differences of reconstructions with and without the platform were smaller than the estimated reconstruction noise.

Fig. 2 shows a portion of an input step recorded by the 4DFRS while Fig. 3 displays the corresponding reconstructed outputs; only four frames of the sequence are displayed, but a common step covers 20 to 40. The time needed to reconstruct each frame is around 1 minute with the current software implementation.

B. Accuracy and Repeatability

We tested the accuracy of the system first by imaging a 30cm × 15cm plane made from rigid wood and painted in matt grey. The dimensions of the plane compare well to the size of the foot. The plane was placed in the workspace in ten different positions and orientations: zero inclination and elevation, zero inclination and 10 cm elevation, $\pm 45°$
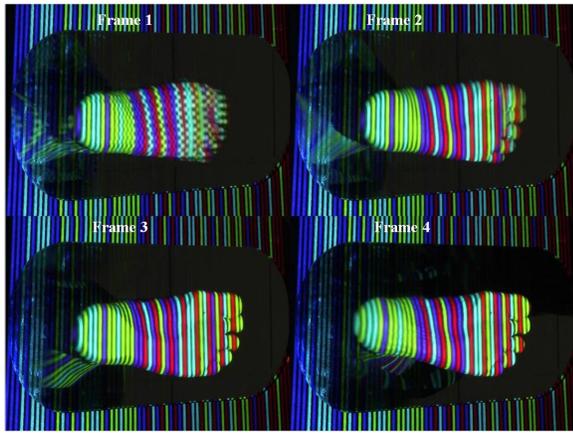
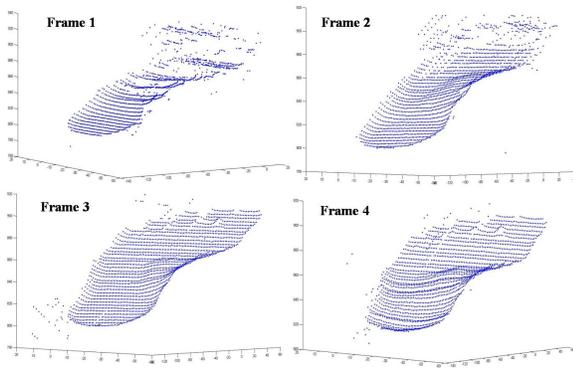Fig. 2. Four frames taken from a video sequence of a step with the CSL pattern projected on top of the foot

Fig. 3. Reconstructed shape (clouds of points) of the foot surface from the frames in Figure 2

Fig. 4. Registered static foot reconstructions using ICP (Blue foot corresponds to day 1 while red to day 2)

from the x-axis, $\pm 45°$ from the y-axis, and $\pm 45°$ from both the x- and y-axis. For each case, a plane was fit to the reconstructed clouds of 3D points. Errors were computed as the euclidean distance of each point from the best-fit plane. The calculated mean absolute errors and standard deviations had a maximum value of 0.5 mm. Two comments must be mentioned about these obtained results. First, they were obtained with static planar surfaces, and must therefore be regarded as an experimental upper limit for the accuracy of the system when used with real feet in static and dynamic conditions. Second, the errors measured are much smaller than variations in foot length meaningful for orthosis design.

Repeatability was assessed by capturing 20 images of a plane. We used two different positions, flat on the ground and slightly elevated, simulating two orientations of the foot surface taking a step We computed reconstructions as analyzed results. The 20 plane images were taken in succession within less than one minute; the same calibration was used throughout. The best-fit plane was estimated for each reconstruction, and mean absolute errors and standard deviations were calculated. The error was again defined as distance from the best-fit plane. The values of data variation obtained were all less than 0.1 mm, negligible compared with the minimum relevant for orthoses design.
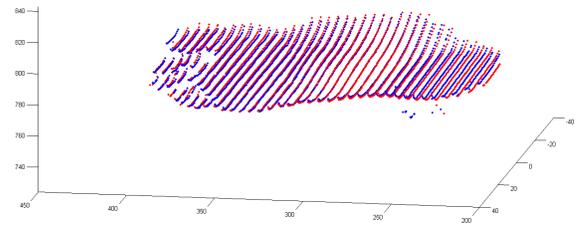
C. Clinical Repeatability

Clinical repeatability is a critical attribute for our application. It is the same repeatability as defined above, but measured with real feet as done in a normal clinical session. The experiments presented in this section were therefore performed using real feet, in both static and dynamic situations.

Twenty-seven subjects were recruited. All the participants were healthy and no prerequisites were imposed on age or gender. For each subject, four static and four dynamic sequences were acquired for both the left and right foot. The subjects were asked to come back one week later to repeat the same measurements. Reconstructions of all the sequences were computed and a comparison performed between 3D foot shapes of each subject from day one and day two.

Figure 4 shows two registered static surfaces taken on two different days (each corresponds to the average of four reconstructions obtained in one trial). Surfaces were registered automatically using the Iterative Closest Point (ICP) algorithm [20], [21].

The average difference between reconstructions of the same foot in the two trials, over the 27 subjects, was found to be 2.440 mm, with standard deviation 2.117 mm.

To compare 4D reconstructions, five frames were selected from each reconstruction sequence, and errors computed for each frame as in the static reconstruction case. The five frames were selected as the key configurations of a foot taking a step: heel strike, arch lowering, stance phase, heel off, toe off [1]. Following the protocol of the experimental case, error statistics were computed for day one and compared to those from day two. An average error was computed for each of the five frames. This average dynamic error corresponds to the average of all the errors obtained from each frame comparison (carried out as in the static case), and was found to be around 3.0 mm.

The errors obtained are very small when compared to possible human errors that appear in current insole casting techniques. These results also compare favourably when compared with the analysis of Maetzler et al. [22] and Ramanathan et al. [23] on the repeatability of different foot measurement systems.

D. Effective Length of the Foot

The effective length of the foot is defined as the length of the foot that is in contact with the ground. This quantity varies in several ways. The *effective dynamic length* is the
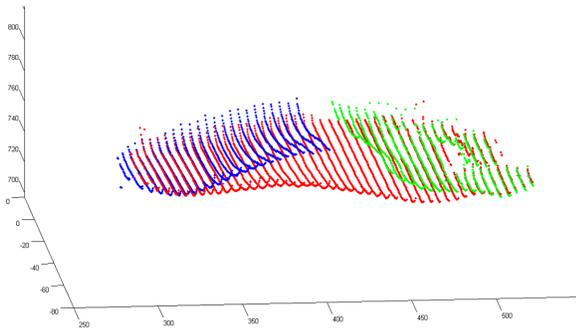
Fig. 5. Dynamic vs. static length. The red reconstruction corresponds to the static foot while the blue represents the heel strike frame and green the toe off frame of the dynamic foot

length of the foot during walking. This is expected to be larger than the static length (Fig. 5) and we used the data from the previous section to test whether there is a significant increase in the length of the foot during walking when compared to the static length.

For the static case, the start point was defined as the closest point to the heel of the foot that is in contact with the ground; the ground was defined as the plane with the highest density of foot points (since the subject is in a standing position), while the end point was chosen to be the farthest point in the fingers that is in contact with the ground. For the dynamic case, the start and end points were selected from the first contact frame (first frame of the foot in contact with the ground) and the last contact one (last frame before toe off) respectively. Notice that the 3D co-ordinates in two different frames of the same sequence can be compared directly as long as the foot does not slip. The start and end points in the dynamic case were selected using the same rule as in the static one, the difference being that the start point was selected from the first contact frame while the last contact frame was used for the end point. Differences between the static and dynamic lengths were computed per subject, per foot and a statistical analysis performed to determine significance.

On average, a length difference of 9 mm was found. Using a significance level of 0.05, all differences between static and dynamic lengths on both days resulted significant in paired t-test, with p-values of all comparisons equal to zero. The data used passed a normality check before the paired t-test was applied.

V. CONCLUSION

We have presented a structured-light prototype system capable of reconstructing the shape of the plantar surface of the foot in motion with accuracy and repeatability sufficient for orthosis design.

This new system is incorporated in a wider research, where dynamic measurements of the foot are obtained and compared to their static counterparts. Substantial difference of around 1 cm was observed between both measurements, suggesting that the use of dynamic foot data provides more accurate results when designing custom made insoles. Our

results suggest that using dynamic reconstruction for clinical purposes can lead to benefits in the way feet are treated. Further studies are needed to quantify such benefits precisely.

REFERENCES

[1] R. J. Abboud, "Relevant foot biomechanics," *Current Orthopaedics*, vol. 16, no. 3, pp. 165–179, 2002.
[2] B. Heiderscheit, J. Hamill, and D. Tiberio, "A biomechanical perspective: Do foot orthoses work?" *British Journal of Sports Medicine*, vol. 35, no. 1, pp. 4–5, 2001.
[3] K. A. Ball and M. J. Afheldt, "Evolution of foot orthotics–part 1: Coherent theory or coherent practice?" *Journal of Manipulative and Physiological Therapeutics*, vol. 25, no. 2, pp. 116–124, 2002.
[4] A. Goodman, "Foot orthoses in sports medicine," *Southern Medical Journal*, vol. 97, no. 9, pp. 867–870, 2004.
[5] C. Payne, "Cost benefit comparison of plaster casts and optical scans of the foot for the manufacture of foot orthoses," *Australasian Journal of Podiatric Medicine*, vol. 41, pp. 29–31, 2007.
[6] Ideas. (2006) Footscanner. [Online]. Available: http://www.ideas.be/
[7] Shoemaster. (2007) Infoot-3d. [Online]. Available: http://www.shoemaster.co.uk/
[8] A. Luximon and R. S. Goonetilleke, "Foot shape modeling," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 46, pp. 304–315, 2004.
[9] A. Luximon, R. S. Goonetilleke, and M. Zhang, "3d foot shape generation from 2d information," *Ergonomics*, vol. 48, pp. 625–641, 2005.
[10] E. Amstutz, T. Teshima, M. Kimura, M. Mochimaru, and H. Saito, "Pca based 3d shape reconstruction of human foot using multiple viewpoint cameras," *International Journal of Automation and Computing*, vol. 5008, pp. 161 – 170, 2008.
[11] J. Wang, H. Saito, M. Kimura, M. Mochimaru, and T. Kanade, "Human foot reconstruction from multiple camera images with foot shape database," *IEICE - Trans. Inf. Syst.*, vol. E89-D, pp. 1732–1742, May 2006. [Online]. Available: http://portal.acm.org/citation.cfm?id=1184858.1185016
[12] T. Coudert, P. Vacher, C. Smits, and M. Van Der Zande, "A method to obtain 3d foot shape deformation during the gait cycle," in *9th International Symposium on the 3D analysis of Human Movement*, June 2006.
[13] J. Pages, J. Salvi, C. Collewet, and J. Forest, "Optimised de bruijn patterns for one-shot shape acquisition," *Image and Vision Computing*, vol. 23, no. 8, pp. 707–720, 2005.
[14] L. Zhang, B. Curless, and S. M. Seitz, "Rapid shape acquisition using color structured light and multi-pass dynamic programming," *Proceedings of the First International Symposium on 3D Data Processing Visualization and Transmission*, 2005.
[15] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, 2003, pp. I/195–I/202.
[16] J. Salvi, J. Pages, and J. Batlle, "Pattern codification strategies in structured light systems," *Pattern Recognition*, vol. 37, no. 4, pp. 827–849, 2004.
[17] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, 2nd ed. Cambridge, UK ; New York: Cambridge University Press, 2003.
[18] J. Salvi, X. Armangue, and J. Batlle, "A comparative review of camera calibrating methods with accuracy evaluation," *Pattern Recognition*, vol. 35, no. 7, pp. 1617–1635, 2002.
[19] T. P. Monks, J. N. Carter, and C. H. Shadle, "Colour-encoded structured light for digitisation of real-time 3d data," *International Conference on Image Processing*, pp. 327–330, 1992.
[20] P. J. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, pp. 239–256, 1992.
[21] E. Trucco, A. Fusiello, and V. Roberto, "Robust motion and correspondences of noisy 3-d point sets with missing data," *Pattern Recognition Letters*, vol. 20, pp. 889–898, 1999.
[22] M. Maetzler, T. Bochdansky, and R. J. Abboud, "Normal pressure values and repeatability of the emed-st2 system," *Gait and Posture*, 2010.
[23] K. P. Ramanathan, A. K., G. P. Arnold, W. Wang, and R. Abboud, "Repeatability of the pedar-x in-shoe pressure system," *Foot and Ankle Surgery*, vol. 16, pp. 70 – 73, 2010.