The virtual human face: Superimposing the simultaneously captured 3D photorealistic skin surface of the face on the untextured skin image of the CBCT scan

K. B. Naudi, R. Benramadan, L. Brocklebank, X. Ju, B. Khambay, A. Ayoub
Biotechnology and Craniofacial Sciences (BACS) Research Group, Glasgow University, College of Medicine, Veterinary and Life Sciences, Glasgow, UK

K. B. Naudi, R. Benramadan, L. Brocklebank, X. Ju, B. Khambay, A. Ayoub: The virtual human face: Superimposing the simultaneously captured 3D photorealistic skin surface of the face on the untextured skin image of the CBCT scan. Int. J. Oral Maxillofac. Surg. 2013; 42: 393–400. © 2012 International Association of Oral and Maxillofacial Surgeons. Published by Elsevier Ltd. All rights reserved.

Abstract. The aim of this study was to evaluate the impact of simultaneous capture of the three-dimensional (3D) surface of the face and cone beam computed tomography (CBCT) scan of the skull on the accuracy of their registration and superimposition. 3D facial images were acquired in 14 patients using the Di3d (Dimensional Imaging, UK) imaging system and i-CAT CBCT scanner. One stereophotogrammetry image was captured at the same time as the CBCT and another 1 h later. The two stereophotographs were individually superimposed over the CBCT using VRmesh. Seven patches were isolated on the final merged surfaces. For the whole face and each individual patch: maximum and minimum range of deviation between surfaces; absolute average distance between surfaces; and standard deviation for the 90th percentile of the distance errors were calculated. The superimposition errors of the whole face for both captures revealed statistically significant differences ($P = 0.00081$). The absolute average distances in both separate and simultaneous captures were 0.47 and 0.27 mm, respectively. The level of superimposition accuracy in patches from separate captures was 0.3–0.9 mm, while that of simultaneous captures was 0.4 mm. Simultaneous capture of Di3d and CBCT images significantly improved the accuracy of superimposition of these image modalities.

Interest in utilizing three-dimensional (3D) images in the planning for orthognathic surgery is increasing because they are considered the ideal methods for representing the face. Creating a precise 3D replica of the head including hard and soft photorealistic tissue structures has been the target of much research. One of the most promising methods is the registration of skin surface images acquired by stereophotogrammetry and cone beam computed tomography (CBCT).

It is generally agreed that creating 2D models is of limited significance, because they are only useful for profile prediction planning. In everyday life patients do not look at themselves in profile. The complexity of some of the suggested registration methods is also considered a serious shortcoming. Relying on laser scanners as a source for the soft tissue data has limitations. The capture is slow, therefore image distortion caused by movement of the subject or change in facial animation is a potential source of errors. In addition, the developed skin surface image lacks a photorealistic appearance and the characteristic surface texture.[1,2]

Stereophotogrammetry, first suggested for use in dentistry by Mannsbach in 1922,[3] makes use of two of more images of an object taken from different viewpoints. A 3D image of the object is built using triangulation to recover the third dimension, thus providing the illusion of depth in the created 3D image.[4] Over the past decades the technique has undergone significant development. The introduction of high resolution digital cameras has allowed the resolution of fine details on the subject's skin surface.[5] Stereophotogrammetry has now developed into a relatively simple, safe, non-invasive, extremely rapid (<1 ms) and highly accurate image capture technique.

CBCT was introduced in 1982 by Robb.[6] It provides high image accuracy with shorter scanning times and lower radiation doses compared to conventional CT.[7] Although CBCT is used in the maxillofacial region primarily to obtain images of the hard tissues of the face, the image editing software used to manipulate the data obtained from the scan is capable of extracting the soft tissue image of the face of the subject. The created image lacks the lifelike photographic texture of the facial soft tissues.

The superimposition of the two images obtained from the above methods would allow the placement of a high resolution 3D facial photograph onto the untextured image of the face obtained from the CBCT image. The difficulty would be ensuring that the facial expression is exactly the same for both image captures. Differences in facial expression could be minimized if the two images were captured at the same time.

The aim of this study is to evaluate the impact of the simultaneous capture of stereophotographs and CBCT images on the accuracy of their registration and superimposition.

Materials and methods

The study was carried out on 14 patients who were being referred for the management of dentofacial problems. Male patients with facial hair were excluded to avoid artefacts in the created image. 3D facial images for these patients were acquired using the Di3d imaging system and i-CAT CBCT scan, which is the normal practice for the authors' orthognathic surgery patients. For each patient, two stereophotogrammetry images were captured, one at the same time as the CBCT scan, this will be referred to as the simultaneous image, and the other delayed image, was captured 30 min later in a separate room. The image for the first stereophotograph was taken while the patient was sitting in the i-CAT scanner just before the CBCT scan was acquired (Fig. 1). Before capturing the images the patients were asked to remove spectacles and jewellery, to keep all hair off the face, to keep both eyes open, bring their teeth in contact and relax their lips.

Stereophotogrammetry

The Di3D imaging system (Dimensional Imaging, Hillington Park, Glasgow, UK) is based on the stereophotogrammetry concept and is able to capture high resolution, full-colour 3D models of the face (180°, ear to ear view). The system consists of two camera stations which are connected to each other. Each station contains a pair of high-resolution (14 Megapixel, 50 mm focal length) digital cameras (Canon (UK) Ltd.). Two white-light studio flash units (Esprit Digital DX1000, Bowens, UK) are placed alongside the camera stations to illuminate the subjects (Fig. 1). The system is calibrated before each use using a fully automated method with a calibration target (Fig. 2).

CBCT

The i-CAT (Imaging Sciences International, Hatfield, USA) is a CBCT imaging tool which is routinely used in many maxillofacial units. Apart from hard tissue information, the image created by the i-CAT can also be manipulated to show the soft tissues of the patient's face. All the patients were scanned using an extended height field of view option (22 cm), 0.4 voxels, with two 20 s scans to capture the complete dataset.

Data processing

DI3Dcapture™ software was used to process the captured stereo pairs of images and create the 3D facial models, as described in Khambay et al.[1] DI3Dview™ software was used to view the created high resolution 3D models and the images were stored as wavefront object files (*.obj).

The CBCT data were imported into Maxilim software (Medicim NV, Mechelen, Belgium) as digital images and communications in medicine (DICOM) files. This allowed manipulation of the image and segmentation of the hard and soft tissues by thresholding. This was done by using the default setting of the Maxilim software which automatically segments the hard and soft tissue surfaces and detects the air tissue boundary interface. The untextured soft tissue surface of the CBCT scans was exported as a stereolithography file (*.stl) (Fig. 3). Maxilim was also used to convert the wavefront object files of the stereophotograph images into stereolithography files to

Fig. 1. The patient in position in the i-CAT machine with the Di3D system positioned in front of the i-CAT ready for image capture.