Using the Microsoft Kinect for Patient Size Estimation and Radiation Dose Normalization: Proof of Concept and Initial Validation

Tessa S. Cook · Gregory Couch · Timothy J. Couch · Woojin Kim · William W. Boonn

Published online: 24 January 2013
© Society for Imaging Informatics in Medicine 2013

Abstract Monitoring patients' imaging-related radiation is currently a hot topic, but there are many obstacles to accurate, patient-specific dose estimation. While some, such as easier access to dose data and parameters, have been overcome, the challenge remains as to how accurately these dose estimates reflect the actual dose received by the patient. The main parameter that is often not considered is patient size. There are many surrogates—weight, body mass index, effective diameter—but none of these truly reflect the three-dimensional “size” of an individual. In this work, we present and evaluate a novel approach to estimating patient volume using the Microsoft Kinect™, a combination RGB camera-infrared depth sensor device. The goal of using this device is to generate a three-dimensional estimate of patient size, in order to more effectively model the dimensions of the anatomy of interest and not only enable better normalization of dose estimates but also promote more patient-specific protocoling of future CT examinations. Preliminary testing and validation of this system reveals good correlation when individuals are standing upright with their arms by their sides, but demonstrates some variation with arm position. Further evaluation and testing is necessary with multiple patient positions and in both adult and pediatric patients. Correlation with other patient size metrics will also be helpful, as the ideal measure of patient “size” may in fact be a combination of existing metrics and newly developed techniques.

Keywords Radiation dose monitoring · Patient size estimation

Introduction

There is growing interest in being able to accurately and routinely monitor patients' exposure to ionizing radiation. Historically, whole-body effective dose has been estimated by multiplying the dose–length product (DLP) by the anatomy-specific conversion factor, or *k* factor, derived from tissue-specific weighting factors determined by the International Commission for Radiological Protection. While deriving effective dose from DLP provides a straightforward, practical estimate of patients' radiation exposure, it does not reflect patient size [1]. Using Monte Carlo simulations of CT scans and subsequent calculation of organ doses, researchers have demonstrated that deriving effective dose from DLP can underestimate the radiation delivered to smaller patients, including children, and overestimate how much radiation is received by larger adults [2, 3].

Pediatric patients are known to be at increased risk for adverse events from overexposure to imaging-related radiation, due to a combination of increased mitotic activity in cells as well as the presence of less attenuating subcutaneous tissue [4]. This knowledge, in combination with the increased utilization of CT in pediatric patients over the past decade, has prompted the formation of the Alliance for Radiation Safety in Pediatric Imaging and the development of the ImageGently® campaign [5]. ImageGently® provides guidance regarding pediatric-specific imaging protocols as well as educational materials for physicians and patients' families.

As McCollough et al. have described, it is critical to understand that the volume CT dose index (CTDI_{vol}) and related dose indices do not represent the radiation dose

T. S. Cook (✉) · W. Kim · W. W. Boonn
Department of Radiology, Hospital of the University of Pennsylvania,
1 Silverstein, 3400 Spruce Street,
Philadelphia, PA 19104, USA
e-mail: tessa@alumni.upenn.edu

G. Couch · T. J. Couch
Radimetrics, Inc, Toronto, Ontario, Canada

actually received by a patient during a CT examination [1]. This discrepancy has motivated the AAPM to develop correction factors for $CTDI_{vol}$ in pediatric and small adult patients, based on effective patient diameter as well as anterior–posterior and/or lateral diameters [6]. Additional work has been done to normalize for inherent differences in scanner geometries and enable comparisons of dose estimates between scanners; however, these corrections still do not account for patient factors such as size, gender, and body habitus [7, 8]. Recent work using Monte Carlo simulations to calculate organ doses from anthropomorphic phantoms of different sizes has more clearly illustrated how much DLP can vary with patient size [9, 10]. At present, these Monte Carlo-based methods are computationally intensive, rendering it impractical to model every patient individually and necessitating the need for standardized phantoms. However, in the ideal situation, every patient would be modeled individually and accurately to provide the best possible dose estimates; this is an area of ongoing research.

There are many potential estimates of patient size: height, weight, body mass index, effective diameter of the anatomy being imaged, etc. However, all of these are either one- or at most two-dimensional measurements, and none of these metrics effectively captures the size of the entire region of anatomy being imaged. To address this limitation, we present and evaluate a novel approach to estimating patient volume using the Microsoft Kinect™, a combination device consisting of an RGB camera and an infrared depth sensor. The goal of using this device is to generate a three-dimensional estimate of patient size, in order to more effectively model the dimensions of the anatomy of interest and not only enable better normalization of dose estimates but also promote more patient-specific protocoling of future CT examinations.

Materials and Methods

Approval from our Institutional Review Board was obtained to evaluate this novel patient volume estimation system using both healthy volunteers and patients. In this manuscript, we focus on initial testing with healthy volunteers. We use a Microsoft Kinect™ sensor that includes an infrared (IR) camera, an IR projector, and an RGB camera. The IR projector produces a speckle pattern. The distortion that results from the speckle pattern is recorded by the IR camera and converted into a depth map for the scene. This depth map is recorded at up to 30 frames per second. The output of the RGB camera was not used for this work. To derive an estimate of whole-body volume, several steps are necessary. Prior to a measurement, the system is calibrated by recording the background and averaging several frames for the purpose of noise reduction. Then, the subject stands on an

electronic scale. The act of standing on the scale not only records the subject's weight automatically, but also triggers the recording of several depth map frames. These frames are also averaged, and the previous calibration background image is used as a mask to segment and isolate the subject. The segmented depth map is then converted to a point cloud representation, and the body volume of the subject is estimated using a convex hull algorithm.

The depth map is displayed on the screen but does not display or record any identifiable features of the subject. Furthermore, for privacy reasons, it is not saved after he or she leaves the viewing frame (Fig. 1). Volume is calculated in real time while the individual remains in front of the device. The volume estimate is continuously refined and recalculated as long as the individual remains within the camera's field of view.

To validate the feasibility of using this Kinect™ camera–sensor for estimating patient size, we calculated the volume of five adult non-patient volunteers and compared this value to each individual's weight. No identifiable information about the volunteers was stored; age, gender, weight, and volume were recorded anonymously. The effect of position on the volume estimation was also evaluated by estimating each volunteer's whole-body volume with the arms by their sides, extended above the head or crossed over the chest.

We hypothesized that the depth information provided by the Microsoft Kinect™ would provide an accurate estimate of whole-body volume. We assumed the density of the body estimated from this volume calculation would be close to $1,000 \text{ kg/m}^3$, which is the density of water, as the human body is predominantly composed of water. We also

Fig. 1 Screen shot demonstrating the depth “image” and derived skeletonization for an adult, produced by the Kinect-based volume estimation system. The image is not stored after the volume estimate is calculated. No identifiable information about the individual is retained by the system

hypothesized that arm position would cause variations in the volume estimation, specifically, that volume estimates with the arms crossed over the chest would differ from volume estimates when the arms are by the sides or extended above the head but not overlapping with the body.

Results

Weight and volume estimates for a total of five adult non-patient volunteers (two males and three females, 36–45 years old) were compared. The volume estimates were obtained with each individual positioned approximately 2.5 m in front of the Kinect™ camera system, standing on a scale to register their weight. To test the repeatability of the volume estimates, multiple depth images were analyzed for each individual. Repeated volume estimates with the individual in the same position each time were noted to be within one significant digit. Figure 2 demonstrates the depth map produced by the system for four of the volunteers. In this figure, the volunteers are all standing with their arms to the sides. Figure 3 demonstrates the additional arm positions that are evaluated: arms extended outward (left) and arms crossed over the body (right).

Figure 4 compares the volume estimates computed for the three different arm configurations. There is a consistent trend of volume increasing when the arms are extended outward and decreasing when they are crossed in front of the body. The difference between the upright–arms down volume estimate and the upright–arms crossed volume estimate is smaller than the difference between the upright–arms down estimate and the upright–arms extended estimate. This trend also affects the density estimates, though the actual volume of the individual volunteers is not changing.

Density is calculated by dividing each individual's weight by his or her volume. We hypothesized that the density of each individual would be close to $1,000 \text{ kg/m}^3$, or the

Fig. 3 Depth maps illustrating the other two positions analyzed: arms extended outward and arms crossed over the body

density of water, as the human body is primarily composed of water. The average density for the five adult volunteers was 995.2 kg/m^3 with the arms by the sides, 805.1 kg/m^3 with the arms extended outward, and $1,136.8 \text{ kg/m}^3$ with the arms crossed over the chest. When the arms were extended outward, the body appeared to occupy a larger volume, but a higher proportion of that volume was composed of air, resulting in a lower density estimate. Similarly, when the arms were crossed over the body, the body occupied a lower volume but appeared to be composed of solid matter, thus increasing the estimates. Some of this variation may arise from the overall angulation of the body that can occur out of the coronal plane when the arms are positioned differently. Nevertheless, this variation with position motivates a more robust volume estimation that can account for these positional differences.

Discussion

Estimation of patient volume using the depth information produced by Microsoft Kinect™ is demonstrated to be repeatable, as well as correlative with patient weight when

Fig. 2 Sample depth maps for four of the adult volunteers, with their arms by their sides, produced with the Kinect system

Fig. 4 Comparison of volume estimates obtained with the arms in different positions in the five adult volunteers

the density of the body is assumed to be similar to the density of water. However, the whole-body volume estimate can provide additional information about the “thickness” of a patient, which is relevant not only to how much dose the patient receives during an imaging study but also to how much radiation is necessary to produce diagnostic-quality images. While this volume estimate is of the entire body of an individual, it can be matched to the region of the patient being imaged in order to provide a regional size estimate and ultimately more accurate size-corrected radiation dose estimates. In addition, it provides valuable feedback about the relationship between the extremities and the torso, as well as the dimensions of each, which cannot be derived from weight alone and may impact proper positioning of the patient for a complicated examination (e.g., imaging of multiple discontinuous body parts in trauma patients). Though the initial work presented here was conducted in healthy volunteers who were not subsequently undergoing a clinical CT examination, patients evaluated in future work will also have coincident scout and cross-sectional imaging. These additional data will be used to further validate the regional volume estimates performed with the Kinect™-based system.

Not surprisingly, volume estimation varies with patient positioning. When the volunteers are standing upright with their arms by their sides, there is a good correlation between the calculated body density and the density of water, which

is an expected result despite the small sample size. This analysis will be extended to a larger cohort of both patient and non-patient volunteers to verify that it is consistent. The effect of position on volume estimation will be explored in future work, specifically with respect to upright versus supine positioning and the variable positioning of the arms with respect to the torso. To improve the accuracy and consistency of volume estimates across the different positions, perpendicular depth maps may be

Fig. 5 Example of interference that can occur with volume estimation when a “foreign body,” in this case, an office chair (*pink structure* to the right of the volunteer), enters the field of view of the camera

necessary. These could be obtained either with two Kinect™ cameras positioned at a 90° angle to one another or by recording depth maps with the patient in consecutive frontal and lateral positions with respect to a single camera. Yet another possible consideration is incorporation of both anterior and posterior views using a two-camera system, instead of doubling the data currently obtained from the anterior view. Regardless of the solution, this is an important consideration as patients are often positioned on the scanner table with their arms extended overhead, and accurate volume estimation in this position will be critical for the success and utility of our Kinect™-based system.

Additional work is also needed to further explore the differences introduced by patient height and age, and to determine what additional modifications are necessary to accurately estimate volume in these individuals. Volume estimation in pediatric patients may pose additional challenges and needs to be tested thoroughly. The effect of respiratory motion on volume estimates also needs to be evaluated, as these have the potential to make the patient seem artificially larger or smaller. Furthermore, the effect of “foreign bodies” such as additional objects within the field of view, baggy gowns or clothing, as well as the CT scanner table itself needs to be evaluated (Fig. 5).

Correlation with other patient size metrics, such as height and body mass index, will be another interesting area of investigation, as the ideal measure of patient “size” may in fact be a combination of existing metrics and newer approaches such as estimated volume. Comparison of Kinect™-based volume estimates to actual body volumes as computed by air displacement plethysmography could also provide further validation of this technique, as well as a better analysis of the inherent error introduced by clothing or patient gowns. Furthermore, it is plausible that the volume estimates generated by the Kinect™ system could be used to perform more accurate whole-body dosimetry for patients undergoing CT examinations.

A number of challenges remain before this patient volume estimation system could be implemented in the clinical realm. Some of these challenges pertain to the use of a consumer device in the medical setting. The Microsoft Kinect™ has been approved by the FDA for navigation and manipulation of imaging studies being viewed in the operating room [11]. Our proposed application would similarly require specific approval for the intended application. In addition, this approval would be heavily contingent on validation of the volume estimation in both adult and pediatric patients, independent of patient position, patient gowns, or additional devices superimposed on or adjacent to the patient. These potential confounders will have to be evaluated with a much larger cohort of individuals than was used in this proof-of-concept work, and potential modifications to the volume estimation algorithm will almost certainly be required to ensure that the accuracy of the volume estimation is

not affected by materials external to the patient. Since the system could foreseeably be ceiling-mounted above the scanner table for volume estimation before the patient undergoes a CT examination, the accuracy and reliability of the system in this configuration will also need to be evaluated. There may be vendor-specific constraints that also need to be considered for compatibility of the system with different CT scanner models.

Conclusion

As the need for more accurate dose monitoring increases, the best surrogates for patient size are being sought. While weight and body mass index are easily obtained, they do not always provide an accurate estimate of patient size, particularly of the region of the patient being imaged. Using the preliminary results from this novel method for estimating patient size, we can work towards more accurate means of estimating patient volume. These improved estimates of patient size can ultimately be used to correct CT dose estimates to better reflect both regional and global patient body habitus as well as patient positioning within the scanner bore. Ultimately, more accurate dose estimation improves care of our radiology patients by supplementing and improving the information that radiologists and non-radiologist physicians can use to order and protocol imaging studies appropriately.

References

1. McCollough CH, Leng S, Yu L, Cody DD, Boone JM, McNitt-Gray MF: CT dose index and patient dose: they are not the same thing. *Radiology* 259(2):311–316, 2011
2. Brenner DJ, McCollough C, Orton CG: Is it time to retire the CTDI for CT quality assurance and dose optimization? *Med Phys* 33:1189–1191, 2006
3. Hurwitz LM, Yoshizumi TT, Goodman PC, Frush DP, Nguyen G, Toncheva G, Lowry C: Effective dose determination using an anthropomorphic phantom and metal oxide semiconductor field effect transistor technology for clinical adult body multidetector array computed tomography protocols. *J Comput Assist Tomogr* 31(4):544–549, 2007
4. Brenner DJ, Elliston CD, Hall EJ, Berdon WE: Estimated risks of radiation-induced fatal cancer from pediatric CT. *Am J Roentgenol* 176(2):289–296, 2001
5. Goske MJ, Applegate KE, Boylan J, Butler PF, Callahan MJ, Coley BD, Farley S, Frush DP, Hernanz-Schulman M, Jaramillo D: The Image Gently campaign: increasing CT radiation dose awareness through a national education and awareness program. *Pediatr Radiol* 38(3):265–269, 2008
6. American Association of Physicists in Medicine: Size-Specific Dose Estimates (SSDE) in Pediatric and Adult Body CT Examinations (Report #204), College Park, MD, 2011
7. Turner AC, Zankl M, DeMarco JJ, Cagnon CH, Zhang D, Angel E, Cody DD, Stevens DM, McCollough CH, McNitt-Gray MF: The feasibility of a scanner-independent technique to estimate organ

- dose from MDCT scans: using CTDIvol to account for differences between scanners. *Med Phys* 37(4):1816–1825, 2010
8. Huda W, Mettler FA: Volume CT dose index and dose-length product displayed...what good are they? *Radiology* 258:236–242, 2011
 9. Li X, Samei E, Segars WP, Sturgeon GM, Colsher JG, Toncheva G, Yoshizumi TT, Frush DP: Patient-specific radiation dose and cancer risk estimation in CT: part II. Application to patients. *Med Phys* 38:408–419, 2011
 10. Li X, Segars WP, Samei E, Sturgeon GM, Colsher JG, Frush DP: Patient-specific dose estimation for pediatric chest CT. *Med Phys* 35:5821–5828, 2008
 11. “GestSure.” <http://gestsure.com>. Accessed November 4, 2012