

# MODELESS MOTION BINNING IN CONE BEAM CT

*T. Petrie, D. Magee*

Computer Vision group  
School of Computing  
University of Leeds  
Leeds, LS2 9JT, UK

*J. R. Sykes*

Medical Physics and Engineering  
Leeds Teaching Hospitals Trust  
St James's Hospital  
Leeds, LS9 7TF, UK

## ABSTRACT

We present the first results of a new technique to bin cone-beam projections without imposing any motion model. Such a technique is required for studying motion in regions of the body, such as the pelvis, where motion exists and is unpredictable. All motion information is obtained directly from the projections and the binning is performed through an exhaustive search of possible assignments. Simplifying assumptions using a proposed protocol change, coupled with dimensional reduction using Principal Component Analysis, make the method tractable.

*Index Terms*— Cone beam, motion compensation, PCA

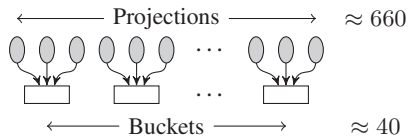
## 1. INTRODUCTION

Cone Beam CT (CBCT) is an imaging modality that reconstructs a 3D volume from multiple 2D kV radiographs taken in a circular path. In image-guided radiotherapy (IGRT), it is used to align the patient with the expected position by rigidly registering the CBCT acquired at the time of delivery with the planning CT used to design the dose delivery. When the registration translation parameters exceed specified clinical tolerances, the treatment couch can be moved precisely using the registration values to reposition the patient such that the treatment volume position is within tolerance. To construct a CBCT on the Electra Synergy machines used at St. James Hospital in Leeds, England, approximately 660 radiographs are taken over the course of two minutes. Once used to verify and possibly correct patient position, the images no longer serve any purpose. Given the x-ray doses the patients receive from this protocol, it is desirable to extract more value from the process. The research discussed in this paper describes an approach that will allow the extraction of information about tumour motion without relying on any kind of motion model. Extracting this motion information has the potential to improve the quality of the CBCT obtained by coupling it with motion compensation techniques (e.g. [1][2][3]), act as the basis for future research on quantifying tumour motion, and provide a mechanism for evaluating the treatment plan margins for motion.

Most of the research in motion compensation of CBCT used in IGRT has been performed in the lung region because of the periodic nature of the motion and the ability to correlate it with various references. Typically, the x-ray images, called projections, are grouped together by phase which is expected to imply similar physical proximity. These groups are called bins. In the simplest case, using only two bins, projections are assigned to either the maximum inhalation phase or the maximum exhalation phase. Usually more than two bins are used. The images are sorted into the bins using a number of techniques: referencing 4DCT scans, measuring the diaphragm movement in the images, measuring respiratory tidal volumes, measuring external surface, etc. Once the projections have been assigned to bins, bin-specific volumes can be reconstructed. A serious problem with binning is the introduction of missing data artefacts and the more bins used, the more incomplete the data used to reconstruct a volume. Likewise, by definition, we are interested in situations where tissue motion has occurred. Such motion introduces its own set of artefacts. Our approach seeks to mitigate the impact both of incomplete data and motion artefacts and provide a basis for binning projections when no motion model and/or no reference information is available. The motion detection, a precursor to compensation, is performed entirely with only the acquired CBCT projections.

## 2. METHOD

Our method is straightforward: we use an exhaustive search to assign each of the projections to a bin. Given two bins and  $n$  projections, and arbitrarily assigning the first projection, we obtain a search space of  $2^{(n-1)}$  possible complete assignments. For the standard protocol, where  $n$  is  $\approx 660$ , this is clearly intractable. By applying one protocol change, some simplifying assumptions, and using established dimensionality reduction techniques, we can reduce the search space and computation time until it becomes possible. The proposed change will have negligible impact on the patient and requires no additional radiation exposure.



**Fig. 1.** First reduction: combining temporally close projections into buckets.

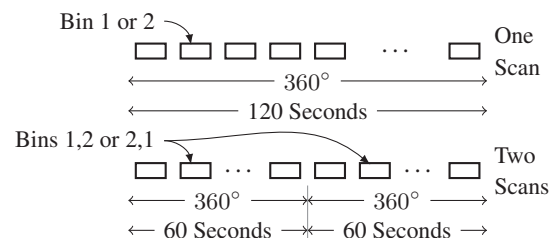
## 2.1. Bucketing

When discussing projections, it is important to highlight the distinction between two conceptually similar attributes: projection *position* and projection *time*. In the standard CBCT protocol, an incremental change in position is coupled to an incremental change in acquisition time. Consequently, in a normal acquisition sequence, which we henceforth call a scan, it is reasonable to presume that sequential projections are adjacent to each other both in terms of acquisition time and angular position. We can use this observation to make our first simplification: adjacent projections in a scan are likely to be assigned to the same bin. We propose partitioning the set of projections into three categories: bin 1, bin 2, and transitional. Bins 1 and 2 hold the projections that we estimate were acquired at two distinct “positions” of the tissue concerned. In the lung region, this would likely correspond to the inhale and exhale states. In the pelvic region, this might represent the before and after states caused by a gastrointestinal movement of some kind. A transitional assignment means a projection’s position information is closer to the mid-point between the positions represented by bins 1 and 2 than to the maxima of the bin positions. We call these groups of projections *buckets*. For breathing scenarios, we can anticipate an average of 8 - 12 breaths per minute. For a two minute scan, this will result in between 32 and 48 buckets to be assigned to either an inhale or exhale state. While this is a significant improvement from  $2^{659}$ , it is still an intractable search space.

## 2.2. Two-pass scanning

Our next reduction involves a protocol change requiring two scans instead of one but with half the projections acquired with each scan and with each scan taking half the time. Such gantry speed up is within the allowable parameters. Consequently, the patient would receive the same dosage in about the same time. The purpose of this protocol change would be to allow us to acquire projections at nearly the same angular position but at different times. We then pair projections for the two scans together by angle similarity. These pairs will, on average, be about .54 degrees apart. Consequently, differences between the projections are caused by very small geometry changes, noise, and motion. When structural contrast is strong enough and where motion has occurred, we see intensity differences between the projection pairs. We use these

pairs both to discover the presence of motion in our region of interest, to compensate for motion outside of our region of interest and to mitigate the missing data problem. By analysing the differences between pairs of projections, we can construct a partition of the angle space. Sequences where the differences between the pairs of projections are small imply that either there was no motion, that we observed the motion at the same phase, or that the motion was hidden by attenuation or noise. In these cases, because we are not imposing any motion model a priori, we simply have no information available and group these together as *similar* projections. In cases where we do see sequences of projections with strong differences, we group these together as *different* projections which will later be assigned, as a group, to one of the two bins. Since we assert that the motion information contained in *similar* projections is insufficient, we average the pairs together and treat them as one projection from the averaged angular position. This prevents the introduction of unwanted differences caused by using the projections independently. We are then left with the task of identifying which bin the projections in the *different* buckets belong to. In the single-scan case, each bucket must be assigned to one of two bins. In the two-scan case, each pair of buckets that exhibit differences in their projections are assigned to either *bin1/bin2* or *bin2/bin1*. Both cases require a binary decision so the number of possibilities, after arbitrarily assigning the first bucket, is  $2^{n-1}$  for  $n$  buckets. However in the two-scan case a complete set ( $360^\circ$ ) of projections is acquired in half the time of a one-scan set so the number of buckets, per scan, will likewise be halved given the temporal size of a given bucket is the same regardless of the protocol used. This reduces our search space from between  $2^{32}$  and  $2^{48}$  to between  $2^{16}$  and  $2^{24}$ .



**Fig. 2.** Top: standard scan bucketing. Bottom: using two scans and assigning buckets (with sufficient projection differences) in pairs. Both are binary assignments but the two scan protocol reduces the bucket count by half.

## 2.3. Clipping to ROI

In radiotherapy situations, several regions are defined during the planning stage [4]. The planning volume is the region that encapsulates the tissue to be irradiated and contains the actual cancerous tissue and the area it could move within during

treatment. We henceforth refer to this as our region of interest (ROI). Our intent is to identify motion within this ROI and mitigate the effects of artefacts in the ROI caused by motion outside of it. Using the ROI, we can construct 2D and 3D masks. Two-dimensional masks can be generated by constructing synthetically generated projections of a solid ROI. A three-dimensional mask is created directly. We mitigate the effects of artefacts caused by motion external to the ROI by averaging each pair of projections together and using the 2D masks to replace the ROI region with the original data for each projection. Motion artefacts created by motion external to the ROI will now be consistent within the ROI regardless of which sets of projection we reconstruct with. This is important because of our search cost function.

Our search cost function is very simple: an assignment is used to generate two reconstructed volumes. These volumes are then compared and the pairs with the strongest sum of squared difference in the ROI are identified as the ones representing the best binary binning of the projection set. Reconstruction is a non-trivial exercise taking on the order of minutes rather than fractions of a second which our  $2^{24}$  upper limit requires. We deal with this by applying a novel approach to the tradition filtered backprojection reconstruction method. By reconstructing a volume for each projection individually, we can average the volumes together to obtain a normal reconstruction. This is equivalent to the typical backprojection process only we move the multiplying constant normally applied to each individual projection outside of the summation process and apply it retrospectively to the backprojection volume. The backprojection phase of the standard filtered backprojection algorithm ([5]), can be tersely described as,

$$B = \int_s \int_t \int_z \int_0^{2\pi} w(s)p(s, t, z, \beta) d\beta \quad (1)$$

where  $B$  is the reconstruction,  $w$  is a weighting function, and  $p$  is the projection that has been filtered. An implementation of this requires discretising the values of the integrals and can be expressed as:

$$\begin{aligned} B_i &= \sum_x \sum_y \sum_z w(x, y, \beta_i) p(x, y, z, \beta_i) \\ B &= \sum_i^N B_i \frac{1}{N} \\ \frac{1}{N} B &= \sum_i^N B_i \end{aligned} \quad (2)$$

where  $B_i$  is an individual backprojection and  $\beta_i$  is one of  $N$  evenly distributed projections. We see that if we reconstruct each projection individually, then our total reconstruction becomes the average of all our individual backprojections.

## 2.4. Applying PCA to Backprojections

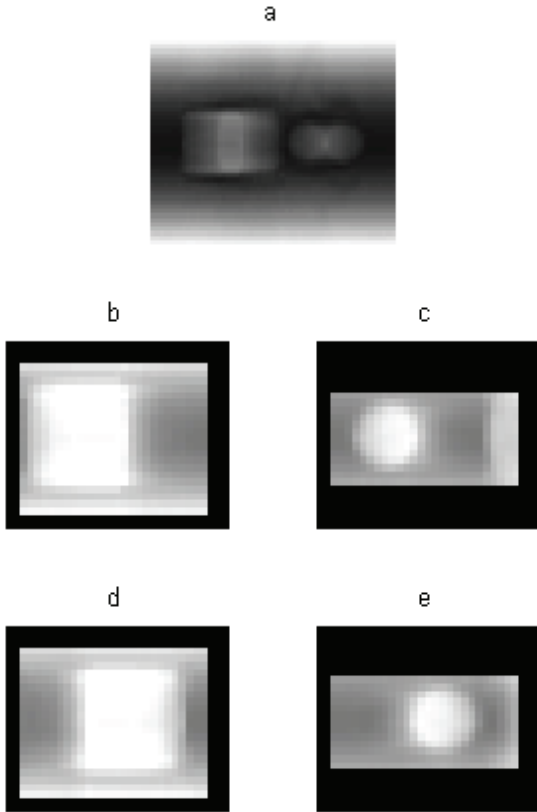
Once we have constructed them, we can apply standard Principal Component Analysis (PCA) to the set of backprojections to perform a dimensional reduction of the data. Given  $m$  projections, PCA will construct an  $m - 1$  dimensional space with our backprojections projected into that space as  $m - 1$  dimensional eigenspace vectors. No information is lost in this transformation as there would be if the number of projections exceeded the number of voxels in our ROI and we were finding a “best” representation in the least-squares sense. Once we have our eigenspace vectors, reconstruction consists of averaging our vectors together, projecting the resultant vector back into our original space, and adding the mean that was subtracted in the initial PCA step. This works because taking the mean in eigenspace is mathematically equivalent to taking the mean in the original space. The other important attribute of vectors in eigenspace is that the relative  $\ell^2$ -norm of the difference between any two vectors is the same as relative difference of the sum of squared differences between the features in our original (voxel) space. In our case, features are the intensity values of the voxels in our reconstructed volumes. This characteristic allows us to evaluate our cost function in the PCA space without requiring a reprojection back into voxel space.

In our process, we have three sets of projections: scan 1, scan 2, and a set from averaging the pairs together. For the suggested protocol, that means approximately 330 projections for each scan and 330 projections for the averaged together set. The dimension of these vectors could be the size of the volumes we are reconstructing. A  $256^3$  length would be common. Using the ROI 3D mask described earlier, we can again substantially reduce this size. Given a  $256^3$  volume and assuming an ROI that is half that dimension, we can compress our feature vectors from 16777216 dimensions (voxels) to 5832 dimensions (voxels). In our test cases, the actual dimension size was much smaller. This compression allows us to perform PCA much more quickly.

Once the projections have been bucketed, compressed, and dimensionally reduced we sequence through the possible permutations of assignments to the partitioning described earlier and evaluate each assignment. Evaluation is performed by summing the vectors from each scan assigned to a given bin, backfilling the missing projections with projections from the average set, averaging them together, and then subtracting the resultant vectors from the two bin reconstructions. This is equivalent to reconstructing the volumes from the projection sets and taking the sum of squared differences of all the voxels in the ROI.

## 3. RESULTS

We show results from a simple virtual phantom with a spherical object and a cube object in sinusoidal motion along the



**Fig. 3.** (a) sample from phantom reconstruction showing cube and sphere in motion (b) sample from bin 1 reconstruction, cube ROI (c) sample from bin 2 reconstruction, cube ROI (d) bin 1 reconstruction, sphere ROI (e) bin 2 reconstruction, sphere ROI. Images windowed to increase contrast.

z-axis with a period of 6.4 seconds. These are embedded in a cylindrical object which is stationary. The sphere object has a radius of 19mm, the cube object is 28mm on an edge and both travel 17.9mm peak-to-peak along the z-axis.

At the top of Figure 3, a slice from the normally reconstructed volume shows the blurred motion of the objects. In the left column, the reconstructions of the cube region of interest show the cube in two states. The same is true for the sphere object in the right column. These reconstructions are from the PCA vectors projected back into voxel space. We have confirmed the fidelity of PCA reconstruction by averaging the principal components from both scans together, projecting back to the original space, and comparing with the tradition reconstruction process. Small observed differences were due to the accumulation of numerical errors. We use 32-bit floating point values to conserve memory usage.

The period choice for the moving objects along with a slightly longer 63 second scan time resulted in partitioning into 20 different buckets. This created a  $2^{19}$  permutation search space. Our code is currently unoptimised, makes no

use of the obvious parallelisation opportunities, and relies on writing intermediate results to disk between stages. Nevertheless, on a machine running a 32-bit OS with 3GB of available RAM and a 2.49GHz CPU, we were able to back-project the projections, run PCA on them, and search the permutation space performing our PCA reconstructions in under 30 minutes.

#### 4. DISCUSSION

We have described and demonstrated a system for extracting motion from slow CBCT imagery without relying on any motion model. This is especially important in regions of the body that exhibit motion but do not have any predictability. Cancers in the pelvic region are particularly susceptible to these conditions. Our research is continuing as we explore how well this technique works on more complex virtual phantoms and on physical phantoms before we propose a clinical trial. We plan to add a second iteration to generate 4-bin assignments and to attempt to generate a simple 4D motion model. We will subsequently attempt to compensate for the motion in the ROI by registering the volumes and generating reprojections from the warped volumes. These reprojections should lead to higher quality cone beam CT images along the lines of [1]. Furthermore, we expect to be able to use the difference information in the paired projections to identify “hot spots” in the volume, generate separate ROIs for each hot spot, and compensate for the motion in each region independently.

#### 5. REFERENCES

- [1] Simon Rit and David Sarrut, “Cone-beam projection of a deformable volume for motion compensated algebraic reconstruction.,” *Conf Proc IEEE Eng Med Biol Soc*, vol. 2007, pp. 6544–6547, 2007.
- [2] J. J. Sonke, L. Zijp, P. Remeijer, and M. van Herk, “Respiratory correlated cone beam ct,” *Medical Physics*, vol. 32, no. 4, pp. 1176–86, 2005.
- [3] T. Li, E. Schreibmann, Y. Yang, and L. Xing, “Motion correction for improved target localization with on-board cone-beam computed tomography,” *Physics in Medicine and Biology*, vol. 51, no. 2, pp. 253–267, 2006.
- [4] International Commission on Radiation Units and Measurements., *Prescribing, recording, and reporting photon beam therapy*, ICRU report. International Commission on Radiation Units and Measurements, Bethesda, MD, 1993.
- [5] L. A. Feldkamp, L. C. Davis, and J. W. Kress, “Practical cone-beam algorithm,” *Journal of the Optical Society of America a-Optics Image Science and Vision*, vol. 1, no. 6, pp. 612–619, 1984.