The PI-methods for Helical Cone-Beam Tomography

Per-Erik Danielsson¹), Maria Magnusson Seger¹), and Henrik Turbell²), 2002
1) Computer Vision Laboratory
Department of Electrical Engineering
Linköping University, SE-581 83 Linköping, Sweden
pedi@isy.liu.se, maria@isy.liu.se
2) SICK | IVP AB
Wallenbergs gata 4, SE-583 35 Linköping, Sweden

Abstract. Next generation helical cone-beam CT will feature pitches around 80 mm. It is predicted that reconstruction algorithms to be used in these machines with still rather modest cone angles may not necessarily be exact, but rather have an emphasis on simplicity and speed. The PI-methods are a family of non-exact algorithms, all of which are based on complete data capture with a detector collimated to the Tam-window followed by rebinning to obliquely parallel ray geometry. The non-exactness is identified as inconsistency in the space invariant one-dimensional ramp-filtering step. It is shown that this inconsistency can be reduced resulting in significant improvement in image quality and increased tolerance for higher pitch and cone angle. A short theoretical background for the PI-methods is given but the algorithms themselves are not given in any detail. A set of experiments on mathematical phantoms illustrate (among other things) how the amount of artefacts grow with increasing cone angles.

1. Introduction

Three-dimensional reconstruction algorithms in general are conveniently divided into exact¹ and non-exact algorithms. The first exact algorithms for helical cone-beam scanning (Tam 1995) were straight-forward extensions of Grangeat’s result (Grangeat 1987), which required that the totality of available projections covered the object in full. However, a practical requirement for any new method is the ability to cope with the long object problem (Danielsson et al 1997). A recent evaluation of three exact algorithms for this purpose has been presented (Sourbelle et al 2001). Even more recent are the methods by Katsevich (2001), Kudo et al (2001) and Hu et al (2001). However, as pointed out by Defrise et al (2001), the numerical complexity of the exact algorithms tends to increase the execution time and call for extra precautions to combat discretization errors and loss of resolution.

The next generation of CT-machines will have moderately high pitches around 80 mm, which means cone angles less than ±2 degrees. Even so, the speed will be impressive. With a gantry that rotates 2.5 r/sec a body section of 20 cm is covered in 2 seconds. Therefore, in the wake of potential implementation problems for the exact methods, non-exact algorithms with inherent simplicity are of considerable interest. The number of such non-exact (approximate) algorithms is steadily increasing (Turbell et al 2000). Among more recent proposals are the one by Stierstorfer et al (2001) where a slice is composed by segmented planes, each one being narrow enough to allow for 2D-reconstruction.

The PI-detector fits exactly the Tam window (Tam 1995), the area between two consecutive turns of the helix. Figure 1 illustrates the ray geometry for this detector in its rebinned and non-rebinned versions. As we show in the next section, the data capture is complete and (almost) non-redundant (Danielsson et al 1997). This is the reason why almost all exact methods are using the PI-detector.

¹ A reconstruction algorithm is exact if any requirement on fidelity to the original can be met, provided that accurate and noise-free projection are data captured with sufficient density along a source trajectory, which meets some specific conditions for completeness.
This report begins with a survey of the basic concepts (such as the PI-line and the full and non-redundant exposure), together with the rational and the theoretical underpinning of the PI-methods. The family of algorithms is PI-ORIGINAL, PI-2D, PI-SLANT, and PI-FAST. None of these are described in any detail, which would require much more space than we are going to allow ourselves in this report. More interesting for the general reader is probably to see experimental proofs that they work and how well they perform under different circumstances. Such experiments with subsequent discussions and conclusions are indeed included.

2. The PI-line

A PI-line is defined to be any line that has two points on the helix less than 360 degrees apart (Danielsson et al 1997). Accordingly, a PI-segment is the segment of a PI-line inside the helix. The PI-reconstruction methods to be described in the next sections are based on the following three observations of the PI-segments.

Observation 1). As the object cylinder moves upwards in the fixed source-detector system in Figure 1a), at the first instant of exposure the PI-segment is a ray-path projecting its content onto a point at the upper edge of the detector. After a 180-degree rotation as seen from a point on the PI-segment itself, this PI-segment becomes a part of a ray-path for the second time. Again a ray projects its content onto a single detector point although at the lower edge of the detector. Hence, all PI-segments are projected over 180 degrees as seen from themselves.
Observation 2). See Figure 2. We study the local Fourier transform of a small sub-volume on a PI-line in the manner of Orlov (1975). Without giving up generality of the following arguments, this sub-volume can be made arbitrarily small so that the distance to the source is so large relative to these dimensions that the divergence of the cone-beam within the small volume is negligible and we may assume that the volume is projected by a parallel beam. The 180-degree exposure can then be portrayed as in Figure 2. The point-shaped object is in the center of a unit sphere. A source path is laid out on this sphere from the entrance angle A to the diametrically opposite exit angle B. We should also imagine that the 3D Fourier transform of the small object is overlaid this space and centered at the origin.

The arbitrary frequency component D is normal to a central plane, which is sectioning the Fourier space as well as the unit sphere onto which we have mapped the source trajectory. By virtue of the Fourier slice theorem, a source position on this plane will generate projections, which contains the frequency component D. The source trajectory from A to B has to cross this plane, say, at C, so that the projection taken from this source position will include D. Therefore, each Fourier component of each small sub-volume on a PI-line is included in the projection data. If A and B are not diametrically opposite there would be room for a central Fourier plane not crossed by the source path. Thus, the 180-degree exposure guarantees complete data capture, which is consistent with Orlov’s theorem (Orlov 1975). Since a great circle from A to B can be drawn through any point on the sphere, the arbitrary source path through C together with one of these great circles can be made to enclose an area as to encompass any given frequency component.

Observation 3). Figure 3 will help us to see that every point inside the helix cylinder is on a PI-line. Select an arbitrary point A. The horizontal plane through A (indicated by a horizontal circle through the helix cylinder) will intersect the helix at B. Let the line BAC rotate around A, while B keeps contact with the helix sliding upwards. During this process, let the other end C keep its contact with the helix cylinder surface. Both C and B are leaving the horizontal plane, but while we force B to move along the helix left- and upwards towards B’, C is moving right- and downwards towards C’. This point on the helix is reached before the line has rotated 180 degrees. The original horizontal line has now become the PI-line B’AC’ to which A belongs. Also, point A cannot belong to any other PI-line. A futile attempt would be to find a second PI-segment as AC” in Figure 3 but in order to land with the other end B” of the line segment on the helix, we have to slide and rotate AC” to the right until it coincides with B’AC’. Hence, each object point belongs to one and only one PI-segment.

The above three observations and conclusions can be summarized in the conjecture that the PI-detector is performing a complete data capture of the object space. This is consistent with the fact that the same data capture has proven to be sufficient and necessary for exact reconstruction (Tam 1995). However, the conjecture also indicates that as soon as the data capture for a small sub-volume is complete it should be possible to reconstruct this sub-volume. In so doing, the long object problem would also be solved. Interestingly, the most recently invented exact method by Katsevich (2001) is indeed able execute its reconstruction without involving projection data from far-away parts of the long object. The same feature also holds for the PI-methods to be presented here.

3. PI-ORIGINAL

In the rebinned geometry Fig. 1b), a set of parallel PI-segments (parallel as seen along the rotation axis) will enter and exit the illuminated area simultaneously. These PI-segments and their object points are residing on a continuous non-planar surface, which we call a PI-surface. As a consequence to the above observation 3), the set of successive PI-surfaces are non-overlapping but mutated with respect to each other, while completely filling the object space.
The discovery of the original PI-method (Danielsson et al 1998) was based on the simple insight that the above rebinned ray geometry guarantees that the backprojection step will deliver the same number of contributions to all object points. No special consideration except simplicity was given to the filtering step. The PI-ORIGINAL algorithm consists of the following four steps.

1) **Pre-weighting** of detector data with cosine of the cone angle
2) **Rebinning** to obliquely parallel projection data on the virtual detector
3) **Rampfiltering** of rebinned data row-by-row
4) **Backprojection** along the original ray paths without magnification factor

Considering the extreme simplicity of the algorithm, we found the results encouraging (see below under 6. Experiments) and a stimulus for further research and improvements. One such development became the \(n\)-PI methods reported by Proksa et al (2000), where \(n\) is any odd number. The original PI-method is the 1-PI method. The 3-PI method utilizes a detector bounded by three consecutive turns; the 5-PI method has a detector bounded by five consecutive turns, etc. It can then be shown (although less trivially than expected) that each object point is indeed exposed during a rotation interval of \(n\pi\). In practice, this family of methods can use (almost) the same physical detector. For the \(n\)-PI case, the patient is just translated \(n\) times slower than for 1-PI, so that the pitch becomes proportional to \(1/n\), all other parameters being constant. With constant photon flux from the source the expected signal-to-noise ratio is proportional to \(\sqrt{n}\). The family of \(n\)-PI methods taken together forms a powerful system able to utilize much higher cone angles than \(\pm 2^0\) by trading speed for artifact and noise reduction.

4. PI-SLANT and PI-2D

**Definitions.** A reconstruction using filtered back-projection is said to be consistent or consistently filtered if the following condition holds. All object points, which exchange energies during the filtering of one set of projection data should also participate in all other filtering events where they participate in the reconstruction process. The more frequent this rule is broken the less is the consistency. Attempts to simplify the cone-beam reconstruction to a 2D-problem typically leads to inconsistency, inconsistent filtering. Another form of consistency can be formulated for the projection back-projection operations. Consistent back-projection uses the same ray-paths as the projection.

Ordinary planar slice-by-slice two-dimensional reconstruction is fully consistent. A typical inconsistency occurs when truly 3D-reconstruction is approximated with 2D-versions. Although we have not (yet) defined a measure of inconsistency, we take it for granted that such a measure can be defined and we allow ourselves to talk about degrees of consistency, minimizing inconsistency and so on. We also take it for granted that there is a strong correlation between inconsistency and non-exactness in the reconstruction. To minimize inconsistency in the original PI-method, we try to localize disjunct 2D sets of object points which are nearly possible to project with a 1D-set of rays throughout the 180 degree rotation. Natural candidates for such sets are the PI-surfaces. The PI-surface is not perfectly planar but slightly saddle-shaped. During the rotation the projection of the points of a PI-surface will be centered along a slanted line, which is the projection of the mid-ray in the PI-surface.

In Figure 4 we see the projections of the points of a PI-surface (onto the planar virtual detector on the rotation axis) for nine rotation positions in the source rotation angle interval from \(\theta = 0^0\) to \(\theta = 180^0\). The detector has the special coordinates \(t\) (horizontally) and \(s\) (vertically). For the top and bottom rows the ramp-filtering along horizontal rows as in PI-ORIGINAL is correct (consistent) with respect to the PI-surfaces. But for the better part of the exposure interval it is
highly inconsistent. In the five-step PI-SLANT algorithm below, the filtering instead takes place along the slanted mid-lines of the projections.

1) *Pre-weighting* of detector data with cosine of the cone angle
2) *Rebinning* to obliquely parallel projection data
3) *Resampling* (vertically) to sets of slanted lines
4) *Ramp-filtering* along slanted lines
5) *Backprojection* along the original ray paths without magnification factor

Filtering of this kind aiming for maximum consistency was proposed by Larson et al (1998) and Heuscher (1999). However, the filtered data were then back-projected into the 2D-surface instead of into the 3D-volume as in PI-SLANT. This results in inconsistency between the actual ray-paths for projection and back-projection. To find out if such a strategy is viable we designed a variation of PI-SLANT we call PI-2D. Here, the back-projection in the last step means back-projection into the two-dimensional PI-surface even if the original projection rays are far from perfectly embedded in the PI-surface. From the reconstruction experiments to be shown below, we draw the conclusion that 3D back-projection along the original rays as in PI-SLANT is a superior technique. Unfortunately, 3PI- and other $n$-PI-methods are not applicable to slanted ramp-filtering.

5. PI-FAST

Although the ramp-filtering inconsistency is significantly reduced in PI-SLANT compared to PI-ORIGINAL it is not reduced to zero. Full consistency would require two-dimensional filtering; seen as an operation in the 3D projection space, also the method by Katsevich (2001) performs two-dimensional filtering. Further reduction of inconsistency in the PI-methods might still be possible with 1D ramp-filtering only. The projection data collected by the virtual PI-detector form a 3D-data set $p(\theta,s,t)$, a kind of 3D-sinogram. Each object point delivers contributions to $p(\theta,s,t)$ along a 3D-curve, which projected onto the $(\theta,t)$-plane is a perfect sinusoid. For each object point in the reconstructed volume the back-projection we are employing is faithfully accumulating the filtered projection data along the same individual 3D-curves.

Figure 4 is a projection of $p(\theta,s,t)$ along the $\theta$-axis, which illustrates how the projection of the non-planar PI-surface tilts and bends as it moves from entrance to exit. The points of the PI-surface are dispersed in the vertical direction, mostly so for $\theta = 45^0$ and $135^0$. The dispersion makes the ramp-filter exchange signal energy between object points in different PI-surfaces, a sign of remaining inconsistency in PI-SLANT. In the PI-FAST-algorithm we attempt to diminish this interaction by letting the ramp-filtering as well as the Back-Projection (BP) operate in the 3D-projection space $p(\theta,s,t)$.
**Fast back-projection** is a technique, which is simple to understand in the 2D-case, (see for instance Brandt et al 1999). Rather than injecting tiny contributions into all voxels for each ray (ray-or detector-driven BP), or summing along the full length of one sinusoid at a time (voxel-driven BP), fast back-projection accumulates limited sets of projection data in the sinogram along short sinusoid segments. These short segments are called *links*. In the next step we sum these link values (with interpolation) into longer links etc, until we have obtained the sums along all full-length sinusoids of interest, which means the image is done. It can be shown that the complexity of \( O(N^3) \) for traditional back-projection can be reduced to \( O(N^2 \log N) \) with fast back-projection, which renders this method the epithet fast.

Figure 5 shows how fast back-projection makes it possible to postpone the ramp-filtering in 2D FB. In Fig.5a) the circles indicate projection data in the sinogram. The slanted lines indicate seven links of length five. Traditional ramp-filtering with a convolution kernel \( g(t) \) followed by back-projection takes place with the upper formula for link-computation in Figure 5. However, as long as the links are parallel this computation is identical to the formula below where we sum (back-project) first and ramp-filter afterwards. In the 2D-case, there is no point in doing this. The complete set of links in Fig.5b) are all in plane and all projection data should interact with each other in this manner anyway. In the 3D-case, however, the corresponding link structure is three-dimensional. Two different sets of parallel links such as in Figure 5b) might not be in-plane but slightly oblique to each other as they follow the slopes of different PI-surfaces.

![Figure 5. Ramp-filtering followed by back-projection over five projection angles \( \theta_j = \theta_0 + \Delta \theta \) equals back-projection over the same projection angles followed by ramp-filtering](image)

The 3D-link structure for a PI-surface is embedded in the \((\theta, s, t)\)-space and can be precompiled in detail (Turbell 2001). At entrance and exit (top and bottom of Figure 4) the 3D-structure of the projection of all points on a PI-surface is thinned down to sheet, while it has a certain thickness in between. This is where PI-FAST makes a difference and the two formulas in Figure 5 take on different values. This algorithm consists of the following six procedures. The links to be used in PI-FAST are of intermediate length covering a limited portion of the \( \theta \)-interval of \( \pi \) radians.

1) **Pre-weighting** of detector data with the cosine of the cone angle
2) **Rebinning** to obliquely parallel projection data on the virtual detector
   For all PI-surfaces
   3) **Back-projection 1**: Accumulating link values over (typically) 16 angles
   4) **Ramp-filtering** the sets of link values
   5) **Backprojection 2**: Accumulating the final pixel values
   6) **Resampling** the pixel values of the PI-surfaces in the z-direction to the Cartesian grid.
6. Experiments

We conducted six experiments with parameters for geometry, reconstruction, and presentation given by the following table. The detector element size (1.56 mm, also used for the reconstruction grid) is measured on the virtual detector at rotation center (2.4 mm on the physical detector). This resolution is only half of what is customary but sufficient for comparison of artifacts. In Experiment 1, presented in Figures 6 and 7, we apply the above four PI-methods to the Clock and the Shepp-Logan phantoms, respectively. The Clock phantom has some features common with the ribs of the human chest, while the well-known 3D Shepp-Logan phantom bears some similarity to a human head with interior low contrast features. The clock phantom consists of an object cylinder with 0 HU on a background of -1000 HU. Inside are two sets of spheres with 400 HU arranged in a clockwise fashion and gradually offset in the z-direction. For image quality evaluation in Figure 6 we also show the same slice (thickness 1.56 mm) reconstructed with the “golden standard” 2D FB, as well as a PI-ORIGINAL slice reconstructed in full resolution (Exp. 0) and slice thickness 0.78 mm. In Experiments 2 - 5, presented in Figure 8, we employ the Clock phantom only. The purpose is to demonstrate the sensitivity to higher fan and cone beams.

<table>
<thead>
<tr>
<th>Exp. 0</th>
<th>Exp. 1</th>
<th>Exp. 2</th>
<th>Exp. 3</th>
<th>Exp. 4</th>
<th>Exp. 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helix radius R</td>
<td>570 mm</td>
<td>570 mm</td>
<td>760 mm</td>
<td>400 mm</td>
<td>760 mm</td>
</tr>
<tr>
<td>Pitch P</td>
<td>81.25 mm</td>
<td>81.25 mm</td>
<td>187.5 mm</td>
<td>100 mm</td>
<td>381.25 mm</td>
</tr>
<tr>
<td>Field of view</td>
<td>400 mm</td>
<td>400 mm</td>
<td>400 mm</td>
<td>400 mm</td>
<td>400 mm</td>
</tr>
<tr>
<td>Fan angle γ</td>
<td>± 20.54</td>
<td>± 20.54</td>
<td>± 15.26°</td>
<td>± 30°</td>
<td>± 15.26°</td>
</tr>
<tr>
<td>Cone angle κ</td>
<td>± 2.04°</td>
<td>± 2.04°</td>
<td>± 3.51°</td>
<td>± 3.58°</td>
<td>± 7.15°</td>
</tr>
<tr>
<td># views per turn</td>
<td>1024</td>
<td>512</td>
<td>512</td>
<td>512</td>
<td>512</td>
</tr>
<tr>
<td># detector rows</td>
<td>52</td>
<td>26</td>
<td>60</td>
<td>32</td>
<td>122</td>
</tr>
<tr>
<td># elem. per row</td>
<td>511</td>
<td>255</td>
<td>255</td>
<td>255</td>
<td>255</td>
</tr>
<tr>
<td>Element size</td>
<td>0.78² mm²</td>
<td>1.56² mm²</td>
<td>1.56² mm²</td>
<td>1.56² mm²</td>
<td>1.56² mm²</td>
</tr>
<tr>
<td>Voxel size</td>
<td>0.78³ mm³</td>
<td>1.56³ mm³</td>
<td>1.56³ mm³</td>
<td>1.56³ mm³</td>
<td>1.56³ mm³</td>
</tr>
<tr>
<td>Slice thickness</td>
<td>0.78 mm</td>
<td>1.56 mm</td>
<td>1.56 mm</td>
<td>1.56 mm</td>
<td>1.56 mm</td>
</tr>
<tr>
<td>Reconstr. grid</td>
<td>512x512</td>
<td>256x256</td>
<td>256x256</td>
<td>256x256</td>
<td>256x256</td>
</tr>
</tbody>
</table>

Figure 6. Clockwise from upper left: 2D filtered back-projection, PI-ORIGINAL (Exp.1), PI-ORIGINAL with full resolution (Exp. 0), PI-FAST (Exp.1), PI-SLANT(Exp.1), PI-2D(exp.1). Gray-level interval ± 50 HU.
Fig. 7. Exp.1 with PI-ORIGINAL, PI-2D, PI-SLANT, PI-FAST. Upper row: xy-slices. Lower row: xz-slices and at far right: No artefacts for truncated long objects (holds for all PI-methods). Grey-level interval: \( \pm 20 \text{ HU} \)

Fig. 8. From top: \((\gamma = 15^0, \kappa = 3.5^0), (\gamma = 30^0, \kappa = 3.5^0), (\gamma = 15^0, \kappa = 7^0), (\gamma = 30^0, \kappa = 7^0)\)
From left: ORIGINAL, 2D, SLANT, FAST. Grey-level interval: \( \pm 50 \text{ HU} \)
7. Discussion and Conclusions

No cone-beam reconstruction can be expected to get better image quality than the 2D reconstructed result from parallel projections shown in the upper left of Figure 6. PI-original contains some streak artifacts and some shadows below some spheres. Still, the image quality is vastly better than PI-2D. We attribute the special type of strong “skewing” artifacts in the PI-2D to the inconsistent back-projection geometry. PI-SLANT, and even more so PI-FAST, have less shadow and streak artefacts than PI-ORIGINAL.

In Exp.0, we used “full resolution” projection data and applied the PI-ORIGINAL-reconstruction algorithm with “full grid resolution” in the reconstructed image with the result at top right in Fig. 6. The full detector resolution at the virtual detector is $0.78^2 \text{mm}^2$, which increases to $1.2^2 \text{mm}^2$ on a physical source-centered detector array at the distance 870 mm.

The size of the cross-cut disks in Figure 6 seems to match the golden standard for all methods except in two cases. The large sphere at 10 o’clock and the small sphere at 5 o’clock are sliced just at the top and in these cases the disks become enlarged in most of the 3D-reconstructed images. The reason is that all 3D-methods involve vertical interpolation on the detector, which creates extra smoothing of edges. In the full resolution PI-ORIGINAL the smoothing/enlargement is proportionally smaller. This seems to be the most significant difference between the two images at top center and right in Fig. 6.

In spite of the narrow gray-level window of $\pm 20 \text{HU}$ in Figure 7, all methods perform remarkably well for the Shepp-Logan phantom, which is known to be sensitive to soft tissue errors.

The geometry in the experiments 2, 3, 4, and 5 (Figure 8) are chosen to find out how gracefully the images reconstructed by four PI-methods are degrading for increasing cone and fan angles. Clearly the image quality ranking from worse to best is here PI-2D, PI-ORIGINAL, PI-SLANT and PI-FAST. The PI-2D-method is doing quite badly for these more demanding geometries. PI-ORIGINAL suffers from shadow artifacts. Some shadows are also visible in PI-SLANT and PI-FAST, although to a much lesser extent. The two upper rows (same cone angle, different fan angles) are almost identical and so are the two lower ones. We conclude that the quality (freedom from artifacts) hinges primarily on the cone-beam angle and to a much lesser extent on the fan angle. This effect is interesting and far from evident, since the PI-surfaces are quite planar for small fan angles, which could be expected to pay off in terms of less artifacts for non-exact methods.

For which cone angle ranges are these methods practically useful? Clearly, precise image quality measures and much more exhaustive experiments are required to answer this question. However, we believe that the strong artifacts for PI-2D in Figure 6 (cone angle $\pm 2^0$ only) makes these methods less useful. We believe this forecast also holds for related methods using nutated surfaces or planes as (Heuscher 1999) and (Larson et al 1998) as well as for all SSRB-techniques (Noo et al 1998) due to inconsistency between projection and back-projection geometries. PI-ORIGINAL should give better images and is probably simpler to implement than these 2D-back-projection methods. PI-SLANT is not quite as simple but delivers somewhat better images. Hence, this method should be practically useful well beyond the cone angle range of $\pm 2^0$, at least for fast scanning and medium image quality imaging. The PI-FAST holds out even better for the rather extreme geometries in Figure 8. Unfortunately, this method is significantly more complicated than the other PI-methods and, in spite of its name, it is also slower.

Acknowledgement

This work was supported by Philips Medical Research, Hamburg, Germany, which is hereby gratefully acknowledged.
References


Heuscher, D. J. (1999), Helical cone beam scans using oblique 2D surface reconstructions, *3D-1999, The fifth int. meeting on fully three-dimensional image reconstruction in radiology and medicine*, Egmond Aan Zee, The Netherlands, June 23-26, pp 204-207


