NOTES ON THE IDT PREPROCESSING AND RGO/DSRI INTERFACE
by I. Lindegren

Introduction
At the Copenhagen meeting (June 12-14) we discussed a number of modifications of the IDT preprocessing and agreed on certain principles and data interfaces. This note is intended to summarize these points, but gives also a proposal for the handling of the coded photon counts.

Decoding of IDT (and SM) counts
The individual samples will be supplied by ESOC in a semi-logarithmic code, so that one byte (8 bits) is always sufficient for a sample. Note that the full range of the byte from 0 to 255 is employed for coding counts in the range 0 ≤ N ≤ 8159. The code consists of a 3-bit exponent (E) and a 5-bit mantissa (M), so that the coded number N' (0 ≤ N' ≤ 255) is

\[ N' = 32E + M \]  
\[ E = \text{int}\left[ \log_2(N+32) - 5 \right] \]  
\[ M = \text{int}\left[ (N+32)2^{-E} - 32 \right] \]

The table below illustrates the coding N → N' and the decoding N' → \( \bar{N} \).

<table>
<thead>
<tr>
<th>N</th>
<th>E</th>
<th>M</th>
<th>N'</th>
<th>( \bar{N} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td>0</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>32-33</td>
<td>1</td>
<td>0</td>
<td>32</td>
<td>32.5</td>
</tr>
<tr>
<td>34-35</td>
<td>1</td>
<td>1</td>
<td>33</td>
<td>34.5</td>
</tr>
<tr>
<td>94-95</td>
<td>1</td>
<td>31</td>
<td>63</td>
<td>94.5</td>
</tr>
<tr>
<td>96-99</td>
<td>2</td>
<td>0</td>
<td>64</td>
<td>97.5</td>
</tr>
<tr>
<td>8032-8159</td>
<td>7</td>
<td>31</td>
<td>255</td>
<td>8095.5</td>
</tr>
</tbody>
</table>

It is seen that N' = N for N ≤ 32, which is true in the vast majority of cases. However for N > 32 the decoding is not trivial, since several different N are encoded as the same N', and must be decoded as a single value \( \bar{N} \). To avoid biases I recommend that the chosen \( \bar{N} \) is the mean of the extreme N-values identically coded. This means that \( \bar{N} \) is given by

\[ \bar{N} = (M+32.5)2^E - 32.5 \]  

which is half an odd integer for N > 31. Decoding is of course effected by means of a look-up table with 256 entries. To avoid floating-point
arithmetic in the binning of samples, I propose that the look-up table gives instead the integer $2\tilde{N}$, and that division by 2 is performed only once per bin while computing the mean count per bin.

**Binning of IDT samples**

My experience with the ML processing of binned data suggests that the errors introduced by the binning itself are below 1 mas with $\geq 45$ bins. This is true also when the phase at the centre of each bin is used to compute the trigonometric factors. However, all these experiments (made on run 2 of the ESTEC data of April 1985) use simulations in which the true scan velocity equals the nominal (168.75 "s). It should be verified that there are no serious resonance phenomena relating to the observing strategy and number of bins, if the true velocity is within 2% of the nominal. With this reservation, I think we can agree on using a binning algorithm with $\approx 45$ bins.

**Definition of reference phase**

The reference phase of IDT sample $\lambda$, $H_\lambda$, is the angle used in the Fourier decomposition of the counts ($N_\lambda$),

$$N_\lambda \sim b_1 + b_2 \cos H_\lambda + b_3 \sin H_\lambda + b_4 \cos 2H_\lambda + b_5 \sin 2H_\lambda$$

(3)

In the process definition (NDAC/LO/005) the reference phase was defined in terms of the reference grid coordinate ($G_\lambda$) computed from available data on the star's position, the attitude, and the field distortions. From the ML estimation of $b_1$ to $b_5$, we get then the phase correction $\beta_3 = \text{ATAN2}(-b_3, b_2)$. The actual phase at frame mid-time is then obtained by adding this correction $\beta_3$ to the reference phase at frame mid-time.

I have later realized that this definition and procedure is unnecessarily complicated, and I propose that the following definition is used instead.

The reference phase is by definition = 0 at frame mid-time (excluding the medium-scale grid distortion, MSD). Let $T_\lambda$ be the frame mid-time and $T_\lambda$ the mid-time of sample $\lambda$. Neglecting any non-linear variation of the smoothed grid coordinate $G$ (i.e., without MSD) in the frame, the reference phase is then simply

$$H_\lambda = 2\pi(T_\lambda - T_\lambda) \frac{d}{G} + \text{MSD}(w, z)$$

(4)

where $G$ is the time derivative of the smoothed grid coordinate. If the samples are numbered $\lambda = 1$ to 2560 in the frame, then $T_\lambda - T_\lambda = (\lambda - 1280.5)/1200$ s. $\beta_3$ is then the desired signal phase at frame mid-time.

With this definition there is no need to compute the grid coordinate very accurately, only the time derivative. Thus a number of effects such as proper motion and parallax can be entirely neglected, and a simplified procedure used for the aberration. (The aberration is needed only to get the field coordinates $w, z$ close enough for the MSD correction.)

**Calculation of $G$ in (4)**

The derivative is obtained by numerical differentiation between two points in the frame, e.g. beginning of the 69th and 253rd slots (these are close to the Gaussian abscissae for the 2-point quadrature formula). Thus we need to compute $G$ at these two points. The procedure could be as follows.
(A) From the analytical attitude representation \( \mathbf{v}(T), \xi(T), \Omega(T) \), compute the heliotropic angles at the required \( T \). Then transform to the equatorial components of the unit vectors \( \mathbf{x}, \mathbf{y}, \mathbf{z} \) (e.g. using my subroutine attit).

(B) From the star catalogue we have the equatorial unit vector \( \mathbf{u} \) for the mean coordinate direction (constant throughout the mission). Add the aberration vector to get an approximate "apparent" direction:

\[
\mathbf{u} = \mathbf{u} + \mathbf{u}
\]  

(5)

Here \( \mathbf{u} \) is the mean barycentric velocity of the Earth, divided by the velocity of light. This vector can be regarded as constant for a period of one day or so. (Note: It does not matter for the further calculations that \( \mathbf{u} \) is not a unit vector.)

(C) The field coordinates \((w, z)\) and field index \((f)\) are then as usual

\[
f = \text{sign}(y')u
\]  \hspace{1cm} (6a)

\[
w = y'u' \cos \frac{1}{2} \gamma - x'u' \sin \frac{1}{2} \gamma
\]  \hspace{1cm} (6b)

\[
z = z'u'
\]

(D) The full expression for the smoothed grid coordinate \( \bar{G} \) as function of \((w, z, f)\) is

\[
\bar{G} = f_{00} + (g_{10} + fh_{10})w + (g_{01} + fh_{01})z +
\]

\[
+ (g_{20} + fh_{20})w^2 + (g_{11} + fh_{11})wz + (g_{02} + fh_{02})z^2 +
\]

\[
+ (B-V - 0.5) \left[ fh'_{00} + (g'_{10} + fh'_{10})w + (g'_{01} + fh'_{01})z \right]
\]  \hspace{1cm} (7)

but for the present purpose it is probably sufficient to take

\[
\bar{G} = g_{10}w
\]  \hspace{1cm} (7')

The derivative \( \bar{G} \) is obtained by numerical differentiation, as mentioned before, while the field coordinates \((w, z)\) for use in (4) are interpolated to each slot \((1/320 \text{ frame})\) from the same two points. We must assume that the MSD is constant in a slot. The MSD is probably a look-up table with \(168 \times 46\) entries corresponding to the scanfields.

Frame data transferred from RGO to DSRI

The following data need to be transferred on a frame-by-frame basis (items marked with an asterisk * are only required for photometry):

(A) Frame header:

- frame mid-time (TDT from 1988 Jan 1.0) [\(\mu\text{s}\)]
- flag for jet actuation during the frame
- number of stars observed in frame \((0 \leq N \leq 10)\)
- heliotropic angles at frame mid-time: \(\nu, \xi, \Omega\)
(B) **Per observed star (object):**

- object ID
- flag (range 0 to 127) for multiplicity, minor planet etc (TBD)
- observed \( \hat{\chi} \) (only fractional part needed), unit 0.0001, range 0 to 1
- m.e. of \( \hat{\gamma} \), unit 0.0001, range 0 to \( \sim \)0.3
- \( \Delta G = G^* - \hat{\gamma} \), unit 0.0001, range \( \pm \)0.3
- \( \Delta \chi^2 \) (3-parameter fit minus 5-parameter fit), unit 0.01, range 0 to 25
- NINT(mean slot/2) (range 0 to 160)

* - field index
* - w (range 0 to 167) [scanfield index]
* - z (range 0 to 45) [""]
* - \( \hat{\beta}_1 \) (unit 0.001, max \( \sim \)8000) [counts/sample]
* - \( \hat{\beta}_2 \) [""]
* - nint(nacc/20) (range 0 to 128)

The exact format (coding of all quantities as non-negative integers) will be worked out by C Petersen. A few remaining problems:

- coding of photometric parameters \( \hat{\beta}_1, \hat{\beta}_2 \) so as to avoid an excessive number of digits on bright stars. Perhaps a non-linear transformation like \( \ln(1+\beta) \) (unit 0.001, range 0 to 10) could be used.

- object numbering for minor planets (new ID for each new frame?)

- definition of the flag for multiplicity etc. Perhaps as follows:

  flag = 0  - object accepted for OTF calibration (implies a 'single' star and reasonable fit)
  1 - 63  - minor planet (there will be at most 63 planets).
            Implies that the object was not used for OTF cal.
  64 - 127 - star rejected for various reasons (up to 6 binary flags possible)