ASTRONOMICAL DATA ANALYSIS SOFTWARE AND SYSTEMS XIV
ASP Conference Series, Vol. 347, 2005
P. L. Shopbell, M. C. Britton, and R. Ebert, eds.

Improved Reduction Algorithm for ISOPHOT-S Chopped Observations

Bernhard Schulz

California Inst. of Technology / IPAC, MC 100-22, Pasadena, CA 91125, USA

Abstract. The astronomical data gathered by the extrinsic photoconductors on-board the ISO satellite were greatly affected by the high energy radiation environment in space. In addition so called transients with long time constants make automated pipeline processing of the data a challenge. ISOPHOT-S is the high sensitivity spectrometer of the ISOPHOT instrument, featuring two 64 element Si:Ga arrays for the wavelength range of 2.5 to 11.6 microns. It collected a wealth of almost 300 observations of extragalactic nuclei at a spectral resolution of ~ 100. Most of those observations were performed using the chopped observing mode, justifying a dedicated effort to revise and improve the automatic processing techniques in order to derive a homogeneously reduced dataset with realistic uncertainties. The improvements compared to the original standard pipeline processing include techniques like ramp subdivision, smoothed sigma kappa deglitching, spike filtering and accounting for a non-Gaussian signal distribution due to the glitch residuum.

1. Introduction

ISOPHOT-S is a low resolution, high sensitivity infrared spectrometer with fixed gratings and a $24''$x $24''$aperture. It was part of ISOPHOT on board of ESA's Infrared Space Observatory (ISO), which was launched Nov 1995 and operated until May 1998[1]. The wavelength coverage comprised two ranges, 2.49 - 4.90 μm (SS), and 5.86 - 11.65 μm (SL) in two 64 pixel Si:Ga arrays (Lemke et al. 1996; Laureijs et al. 2003; Acosta-Pulido & Abraham 2003). The small detector currents were integrated in cold amplifiers, and appear in the downlinked data as integration ramps that are sampled in equidistant time intervals (Figure 1). Observations were executed in terms of Astronomical Observing Templates (AOTs), which in the case of PHT-S always started with a 32 sec dark measurement. This procedure helped to mitigate memory effects of the detectors by starting each observation under the same conditions. The following sky measurement was conducted either in staring mode, chopped mode or raster mode. This paper will concentrate on the data reduction techniques for the chopped mode measurements between a source and one or two background positions, being the instrument's most important mode for observations of faint extragalactic sources.

[1] `http://www.iso.vilspa.esa.es`

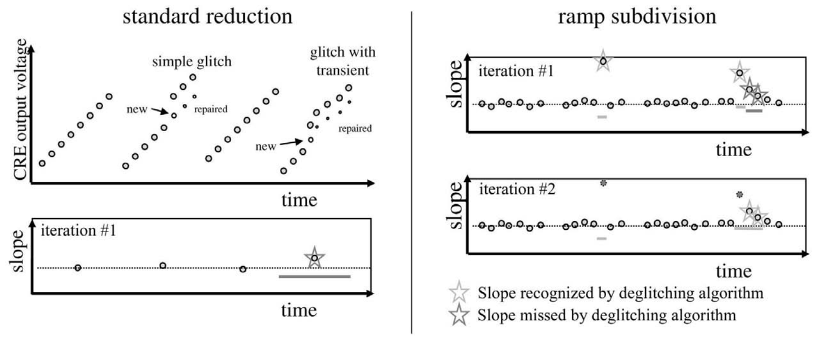

Figure 1. Upper left: Schematic integration ramps affected by a glitch and a spike with transient. Lower left: Effect of old reduction scheme suffering from spike "repair". Right: New scheme using slopes from subdivided ramps. Less integration time is lost by eliminated glitches.

2. The Detector Signal

The ISOPHOT-S data comes in equidistant samples of a slowly rising voltage that is reset in regular intervals. Ideally the photocurrent is determined simply by fitting the slope. Unfortunately high energy particle hits (glitches) create breaks in the integration ramps, and in a number of cases change the responsivity of the detector element for up to minutes (spike). The frequency of time intervals between glitches with a given duration falls exponentially towards longer intervals. Half of the events are less than 8.25 sec apart and most glitches are $\tilde{1}.5$ sec apart. Therefore it is almost impossible to find an integration ramp, which typically lasts 16 to 32 sec for this AOT, that is not broken by glitches.

The standard pipeline processing (OLP) applied the so called two threshold deglitching described in Laureijs et al. (2003). To avoid ramp fragments of different sizes, the algorithm corrects the remaining valid datapoints after a glitch using the estimated slope and fits an overall slope to the entire ramp. This effective insertion of artificial datapoints leaves the calculation of uncertainties less reliable and may introduce larger errors when many readouts in a ramp are flagged after a spike event.

3. Ramp Subdivision

To avoid the statistical problems associated with either artificial datapoints or different size ramp fragments, in this approach integration ramps are broken up into small subramps of equal duration (1 sec). This interval is matched to the distribution of time intervals between glitches as mentioned above.

Before slope fitting no effort is made to "repair" ramps that are obviously glitch affected. This has the advantage that, although the resulting slope distribution is wider due to the smaller effective integration time, the glitch affected subramps become much more prominent outliers in the distribution that can be easily removed by a Sigma-Kappa deglitching algorithm as used by OLP or PIA (Gabriel 1997). If a subramp is removed, a much smaller amount of integration time is sacrificed as if an entire integration ramp of 32 sec must be eliminated.

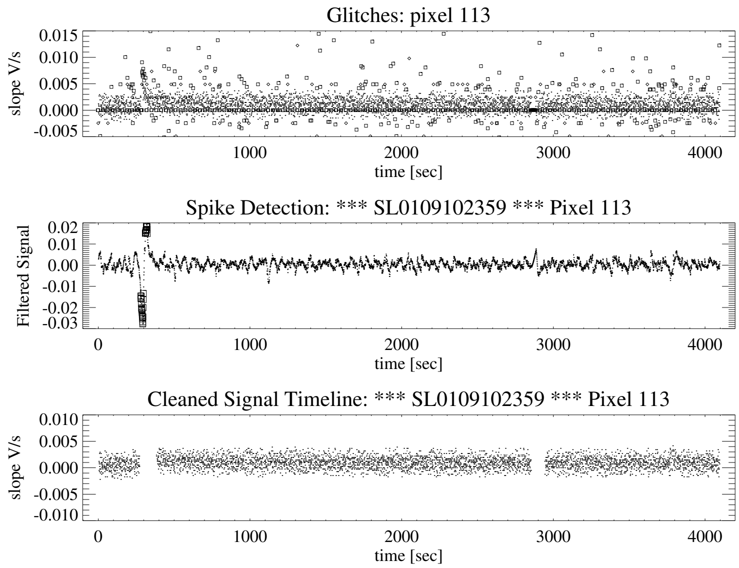

Figure 2. (Upper: Slopes with glitches. Middle: correlation function indicating spike. The second spike is detected by method a). Lower: Cleaned signal.

Without "repairs" the dataset remains "cleaner" and more homogeneous in a statistical sense, since all ramps are of the same size. A drawback of the shorter subramps is the magnification of discrete slope levels due to the digital representation of the data. Therefore the signal determination can not use medians as discussed below.

4. Deglitching

The deglitching algorithm used here differs from the one used in the standard processing, in that it first establishes a signal baseline by smoothing the data with a 60 sec boxcar, calculates a standard deviation σ from this baseline, and discards all points more distant than $\pm 2.5 \times \sigma$. This procedure is repeated until no slopes drop out anymore. The baseline has the advantage of better following longterm drifts that are common in these detectors. This reduction step removes all small energy glitches, however does not catch spike events, which render a longer continuous interval of slopes useless (Figure 2).

The observations described here were made by switching continuously between source and background (chopping) with a minimum of 4 subsequent integration ramps on each position and at least 2 chopper cycles. The background position is located either on one side (rectangular mode) or on opposite sides of the source (triangular mode). Most of the time the signal difference is much smaller than the noise, however it can still broaden the distribution and lead to a systematic lower signal difference due to undetected glitches in the valleys between on-source integrations. In a second phase, a provisionally calculated average difference between on- and off-source pointing is temporarily removed. This enables another deglitching pass to find missed glitches in the valleys between chopper plateaux.

5. Despiking

Strong events (spikes) that change the detector responsivity for up to a few minutes, so far had to be removed manually. Here we describe a new algorithm, that detects and removes spikes automatically. To avoid hiding the spike signature, slopes within 5 σ that were flagged as glitches before, are temporarily included back into the analysis. Two indicators have been found to work most successfully: a) If the ramps were saturated for at least 8 sec, indicated by an uninterrupted sequence of readouts flagged bad and set to zero, the following 90 sec are flagged. b) The slope signal is correlated with a prototype exponential decay signal with a time constant of 90 sec, and the smoothed correlation function is searched for extreme outliers. If at least 5 instances of the correlation function within 90 sec exceed 3 times its standard deviation , all slopes within -9 and +63 sec from this point in time are discarded. See Figure 2.

6. Signal Determination

Once the data is cleaned from glitches and spikes, the on-source and off-source signals must be determined. Sometimes the detector signal shows strong drifts, typically after previous bright illumination. The resulting baseline is fitted with a simple exponential drift model (Huth & Schulz 1998) and removed before determination of the signal. Since the energy distribution of glitches drops continuously to zero, smaller glitches can not be separated from other statistical noise. Therefore and due to residual transients in the data, the slope distribution is not symmetric and shows a small tail towards higher values. The median is closer to the real peak of the distribution, but suffers from the discrete nature of the data (Schulz et al. 2002), since the median can be only an existing value. The best results providing the most reliable signals and uncertainties are obtained by automatically creating separate histograms for on-source and off-source slopes and fitting Gaussians. The difference of their center positions determines the source signal. The uncertainty of each signal is given as the standard deviation of the fitted Gaussian, divided by the square root of the total number of valid slopes.

7. Flux Calibration

The transformation of signals into point source fluxes is still done using the signal dependent relation given by Acosta-Pulido & Abraham (2003). The results are already quite good, however another improvement may be obtained by subjecting the calibration stars to the same processing, to remove any algorithm dependent systematic effects from the flux calibration. This step is still under way at the time of writing.

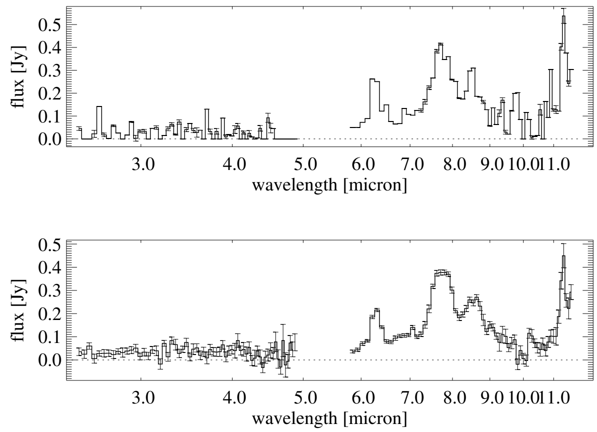

Figure 3. Example of a spectrum taken from the archive (top) and the same data reduced automatically with the described method (bottom).

8. Comparison

In Figure 3 we compare spectra derived from the same raw data set that were obtained by ISO observing the starburst galaxy NGC701. The upper one is a standard pipeline product as found in the ISO archive. The lower is produced automatically using the methods described above. The fluxes of the pipeline product are strongly affected by undetected glitches, and the uncertainties appear unrealistically small. The lower spectrum is less noisy and uncertainties are more consistent with the pixel to pixel noise. The subdivision technique was already used to reduce the Seyfert spectra presented by Clavel et al. (2000); however, the despiking was only automated recently, allowing fully automatic reduction of all chopped ISOPHOT-S observations in this way.

References

Acosta-Pulido, J. & Abraham, P., 2003, ESA SP-481

Clavel, J. et al. 2000, A&A, 357, 839-849

Gabriel, C., Acosta-Pulido, J., Heinrichsen, I., Skaley, D., Morris, H.,& Tai, W.M. 1997, in ASP Conf. Ser., Vol. 125, ADASS VI, ed. G. Hunt & H. E. Payne (San Francisco: ASP), 108

Huth, S. & Schulz, B. 1998, in ASP Conf. Ser., Vol. 145, ADASS VII, ed. R. Albrecht, R. N. Hook, & H. A. Bushouse (San Francisco: ASP), 212

Laureijs, R.J., Klaas, U., Richards, P.J.,Schulz, B.,& Abraham, P. 2003, ESA SP-1262

Lemke, D. et al. 1996, A&A, 315, L64-L70

Schulz, B. et al. 2002, A&A, 381, 1110-1130