Improving the Angular Resolution of Coded-Mask Telescopes by Direct Demodulation *

Zong-Jun Shen and Jian-Feng Zhou

Department of Engineering Physics and Center for Astrophysics, Tsinghua University, Beijing 100084, China; *shenzj98@mails.tsinghua.edu.cn*

Abstract We develop a new procedure to improve the angular resolution of coded-mask telescopes by the Direct Demodulation Method (DDM). DDM has been applied to both real and simulated data of INTEGRAL/IBIS. The angular resolution of IBIS/ISGRI has been improved from about $13'$ to $2'$.

Key words: instrumentation: coded-mask — telescopes: INTEGRAL/IBIS — techniques: high angular resolution — methods: direct demodulation

1 INTRODUCTION

For coded aperture telescopes (Caroli et al. 1987; Goldwurm 1995; Skinner 1995) the source radiation is spatially modulated by a mask with both opaque and transparent elements before being recorded by a position sensitive detector. The mask patterns are designed to allow each source in the field of view to cast a unique shadowgram on the detector in order to avoid ambiguities in the reconstruction of the sky image. This reconstruction (deconvolution) is generally based on a correlation procedure between the recorded image and a decoding array derived from the mask pattern.

Representing the mask with an array $\mathbf{M}$ of 1 (transparent) and 0 (opaque) elements and the detector array with $\mathbf{D}$, the sky image generated by the cross-correlation is:

$$\mathbf{F}_{\mathrm{c}} = \mathbf{D} * \mathbf{G}, \qquad (1)$$

where $*$ represents convolution and $\mathbf{G} = \mathbf{M} - \overline{\mathbf{M}}$ with $\overline{\mathbf{M}}$ the average of $\mathbf{M}$.

It would be ideal to find a mask pattern that makes

$$\mathbf{M} * \mathbf{G} = \delta. \qquad (2)$$

Then,

$$\mathbf{F}_{\mathrm{c}} = \mathbf{D} * \mathbf{G} = \mathbf{F} * \mathbf{M} * \mathbf{G} = \mathbf{F} * \delta = \mathbf{F}, \qquad (3)$$

where $\mathbf{F}$ is the source distribution. Mask patterns with such a property including the uniformly redundant arrays (URA) were found in the 1970 s (Fenimore & Cannon 1978) and were successfully applied in X/γ-ray telescopes onboard several high energy missions (Caroli et al. 1987; Paul et al. 1991). In practice, Modified URA (MURA) (Gottesman & Fenimore 1989) masks are widely used. Despite some fluctuations, random coded masks have this property as well.

It is shown in Equation (3) that the cross-correlation is almost optimal for the image reconstruction of coded mask systems with the property shown by Equation (2). Therefore, the angular resolution of the cross-correlation is treated as the intrinsic angular resolution of the coded mask system.

In practice, Equation (3) is tenable only in the fully coded part of $\mathbf{F}$ of the (M)URA coded mask system. Below, whenever reconstruction of $\mathbf{F}$ is mentioned it is meant that while the whole $\mathbf{F}$ is reconstructed, only the fully coded part is analyzed.

In the theoretical study, The detector elements and the mask elements are usually of the same size : $s_d = s_m$. The angular resolution of such a system is $\arctan(s_m/d)$, d the distance between the detector array and the mask. In practice, however, s_m and s_d are usually not the same. To perform the cross-correlation, the decoding array $\mathbf{G}$ must be resampled at the detector's pixel size. The corresponding angular resolution is $\arctan(\sqrt{s_m^2 + s_d^2}/d)$ (Gros et al. 2003), which is a bit worse than $\arctan(s_m/d)$.

Apart from the cross-correlation method mentioned above, the method of Direct Demodulation (DDM) (Li & Wu 1993, 1994) can also be used to reconstruct the image $\mathbf{F}$. More generally, the modulation process can be represented by the following modulation equation:

$$\mathbf{P} \cdot \mathrm{V}(\mathbf{F}) = \mathrm{V}(\mathbf{D}), \tag{4}$$

where $\mathbf{P}$ is the response matrix of the system, V is an operator and $\mathrm{V}(\mathbf{F})$ the vector form of $\mathbf{F}$. The DDM solves directly the modulation equations by iteration subject to give physical constraints. The iteration used in DDM could be Gauss-Seidel iteration, Jacobi iteration or Richardson-Lucy iteration (Richardson 1972; Lucy 1974), etc. DDM can obtain high signal-noise ratio images, with greatly decreased effect of background and noise. Also, it can greatly improve the location precision and raise the angular resolution to even much higher than the intrinsic angular resolution of the instrument. This technique has been applied successfully to imaging analysis of many missions and different kinds of instruments, e.g., EXOSAT/ME (Lu et al. 1996), HEAO1-A4 (Lu et al. 1995), CGRO/COMPTEL (Zhang et al. 1997, 1998), RXTE/ASM (Chen et al. 2000), ROSAT/PSPC (Chen et al. 1997; Lu et al. 2001), and XMM-Newton (Feng et al. 2003).

In Section 2, the accelerated DDM (ADD) (Shen & Zhou 2007) is applied directly to the imaging of coded mask telescopes in which $s_m > s_d$, which improves the angular resolution to $\arctan(s_d/d)$. Then a technique is introduced in Section 3 to achieve much better angular resolutions than $\arctan(s_d/d)$. In Section 4, real and simulated data of INTEGRAL/ISGRI (Ubertini et al. 2003) are used to demonstrate the performance of these techniques. Section 5 presents a discussion.

For the convenience of description, two operators are defined here: ".$\times$" and "./". $\mathbf{C} = \mathbf{A} . \times \mathbf{B}$ means $\mathbf{C}(i,j) = \mathbf{A}(i,j) \times \mathbf{B}(i,j)$ and $\mathbf{C} = \mathbf{A}./\mathbf{B}$ means $\mathbf{C}(i,j) = \mathbf{A}(i,j)/\mathbf{B}(i,j)$ for every i,j.

2 IMPROVING THE ANGULAR RESOLUTION BY DDM

Theoretically, DDM can be used directly to solve Equation (4), but in this case it consumes extremely large amount of computation time and memory space for the storage of $\mathbf{P}$ which contains $N_F \times N_D$ elements, where N_F and N_D are the number of elements of $\mathbf{F}$ and $\mathbf{D}$, respectively. Fortunately, a coded mask system is a shift-invariant system and the modulation Equation (4) can be rewritten into the convolution form:

$$\mathbf{M} * \mathbf{F} = \mathbf{D}. \tag{5}$$

Consequently we can directly use the ADD to do the DDM imaging efficiently. ADD can accelerate the DDM imaging process by hundreds of times. $\mathbf{M}$, which has fewer than N_F elements, is stored instead of $\mathbf{P}$. This greatly deceases the computation time and memory space. Explicitly, the process is as follows:

1. Estimate the background $\mathbf{B}$ of the image; cf. Li & Wu (1993, 1994).
2. Let
$$\mathbf{S} = \mathbf{W} * \mathbf{ones}, \tag{6}$$
   where $\mathbf{W}$ is obtained by rotating $\mathbf{M}$ by $180°$, $\mathbf{ones}$ is a matrix whose elements are all ones and has the same size as $\mathbf{D}$.
3. Let r represent the current iteration step, $r = 0$, and $\mathbf{F}^{(0)}$ to a matrix and all of which elements are ones.
4. Take equations
$$\mathbf{D}^{(r)} = \mathbf{M} * \mathbf{F}^{(r)}, \tag{7}$$
$$\mathbf{T}^{(r)} = \mathbf{D}./\mathbf{D}^{(r)}, \tag{8}$$

b1	b2	b1	b2	b1	b2	b1	b2	
a1	a2	a1	a2	a1	a2	a1	a2	
b1	b2	b1	b2	b1	b2	b1	b2	
a1	a2	a1	a2	a1	a2	a1	a2	
b1	b2	b1	b2	b1	b2	b1	b2	
a1	a2	a1	a2	a1	a2	a1	a2	
b1	b2	b1	b2	b1	b2	b1	b2	
a1	a2	a1	a2	a1	a2	a1	a2	

Fig. 1 Split every pixel into four sub-pixels and label them a1, a2, b1 and b2, respectively.

$$\mathbf{P}_0^{(r)} = \mathbf{W} * \mathbf{T}^{(r)}, \tag{9}$$

$$\mathbf{F}^{(r+1)} = \mathbf{F}^{(r)} \cdot \times \mathbf{P}^{(r)} ./ \mathbf{S}. \tag{10}$$

5. For every j, h, if $\mathbf{F}^{(r+1)}(i, j) < \mathbf{B}(i, j)$, then $\mathbf{F}^{(r+1)}(i, j) = \mathbf{B}(i, j)$.
6. If the criterion to stop the iteration is not satisfied, then $r = r + 1$ and go to 4.
   The criterion used in the simulation study below is:

$$\frac{|\mathbf{F}_{\mathrm{F}}^{(r+1)} - \mathbf{F}_{\mathrm{F}}^{(r)}|}{|\mathbf{F}_{\mathrm{F}}^{(r+1)}|} < \varepsilon, \tag{11}$$

where $\mathbf{F}_{\mathrm{F}}$ means the fully coded part of $\mathbf{F}$, $|\mathbf{F}|$ equals $\sqrt{\sum_{i,j}(\mathbf{F}(i, j))^2}$ (for every i, j) and ε a parameter which is preset according to the requirement of the imaging accuracy.

In the above procedure, the angular resolution is limited by the pixel size of the detector. For a coded mask telescope with $s_{\mathrm{m}} > s_{\mathrm{d}}$, the decoding array will be resampled at the detector's pixel size, and the corresponding pixel size of the image is $\arctan(s_{\mathrm{d}}/d)$, which is the best angular resolution that this procedure can reach. However, this resolution is probably less than the best angular resolution of DDM.

To achieve better angular resolution, smaller pixel size of the image should be used. Furthermore, a new DDM procedure is required to deal with this case.

3 TO ACHIEVE BETTER ANGULAR RESOLUTION

According to the experiences of the former applications of DDM to other instruments, the DDM can achieve a much higher angular resolution than the intrinsic resolution of the instrument. Therefore, the DDM may achieve a better angular resolution than $\arctan(s_{\mathrm{d}}/d)$ in case of coded mask systems. Because the pixel size of the image $\mathbf{F}$ in Equation (5) is $\arctan(s_{\mathrm{d}}/d)$, we have to use smaller image pixel to represent the flux distribution of the region of interest with better angular resolution. However, this brings on the following problem: the shift-invariant property is broken and Equation (5) does not make sense any more. As a result, ADD is no longer applicable here. The rest of this section presents a technique to solve this problem.

First, every pixel of $\mathbf{F}$ in Equation (4) is split into n sub-pixels. In Figure 1, $n = 4$. These sub-pixels form a new image $\mathbf{E}$. A corresponding response matrix $\mathbf{Q}$ exists. Then the modulation equation is:

$$\mathbf{Q} \cdot \mathrm{V}(\mathbf{E}) = \mathbf{D}. \tag{12}$$

Let $\mathbf{E}_{\mathrm{a1}}$ stand for the matrix formed by sub-pixels labelled by a1, and $\mathbf{Q}_{\mathrm{a1}}$ is formed by the columns of $\mathbf{Q}$ corresponding to $\mathbf{E}_{\mathrm{a1}}$. Change the order of the elements in $\mathrm{V}(\mathbf{E})$ and the order of the columns in $\mathbf{Q}$, we obtain

$$\begin{aligned} \mathbf{Q} \cdot \mathrm{V}(\mathbf{E}) &= [\mathbf{Q}_{\mathrm{a1}}, \mathbf{Q}_{\mathrm{a2}}, \mathbf{Q}_{\mathrm{b1}}, \mathbf{Q}_{\mathrm{b2}}] \cdot [\mathrm{V}(\mathbf{E}_{\mathrm{a1}}), \mathrm{V}(\mathbf{E}_{\mathrm{a2}}), \mathrm{V}(\mathbf{E}_{\mathrm{b1}}), \mathrm{V}(\mathbf{E}_{\mathrm{b2}})]^T \\ &= \mathbf{Q}_{\mathrm{a1}} \cdot \mathrm{V}(\mathbf{E}_{\mathrm{a1}}) + \mathbf{Q}_{\mathrm{a2}} \cdot \mathrm{V}(\mathbf{E}_{\mathrm{a2}}) + \mathbf{Q}_{\mathrm{b1}} \cdot \mathrm{V}(\mathbf{E}_{\mathrm{b1}}) + \mathbf{Q}_{\mathrm{b2}} \cdot \mathrm{V}(\mathbf{E}_{\mathrm{b2}}). \end{aligned} \tag{13}$$

Because the distance between two neighboring pixels of $\mathbf{E}_i$ (i=a1, a2, b1, b2) is identical to the distance between two neighboring pixels of $\mathbf{D}$, $\mathbf{Q}_i \cdot \mathrm{V}(\mathbf{E}_i)$ (i=a1, a2, b1, b2) can represent a shift-invariant subsystem. We can find an $\mathbf{M}_i$ which makes $\mathbf{Q}_i \cdot \mathrm{V}(\mathbf{E}_i) = \mathbf{M}_i * \mathbf{E}_i$. Then

$$\mathbf{Q} \cdot \mathrm{V}(\mathbf{E}) = \mathbf{M}_{a1} * \mathbf{E}_{a1} + \mathbf{M}_{a2} * \mathbf{E}_{a2} + \mathbf{M}_{b1} * \mathbf{E}_{b1} + \mathbf{M}_{b2} * \mathbf{E}_{b2} = \mathbf{D}, \tag{14}$$

where $\mathbf{M}_i$ (i=a1, a2, b1, b2) is equal to the shadowgram of a point source at the central pixel of $\mathbf{E}_i$.

Subsequently, Equations (6) to (11) of ADD can be transformed as (here a pixel of $\mathbf{F}$ is split into n sub-pixels):

$$\mathbf{S}_i = \mathbf{W}_i * \mathbf{ones}, \quad i = 1, 2, \cdots, n. \tag{15}$$

where $\mathbf{W}_i$ is generated by rotating $\mathbf{M}_i$ by $180°$, $\mathbf{ones}$ is a matrix with all elements ones and have the same size as $\mathbf{D}$, and

$$\mathbf{D}^{(r)} = \sum_{i=1}^{n} \mathbf{M}_i * \mathbf{E}_i^{(r)}, \tag{16}$$

$$\mathbf{T}^{(r)} = \mathbf{D}./\mathbf{D}^{(r)}, \tag{17}$$

$$\mathbf{U}_i^{(r)} = \mathbf{W}_i * \mathbf{T}^{(r)}, \quad i = 1, 2, \cdots, n, \tag{18}$$

$$\mathbf{E}_i^{(r+1)} = \mathbf{E}_i^{(r)}. \times \mathbf{U}^{(r)}./\mathbf{S}_i, \quad i = 1, 2, \cdots, n. \tag{19}$$

By using the above procedure, the angular resolution of DDM will no longer be restricted by the pixel size.

4 APPLICATION TO THE INTEGRAL/ISGRI DATA

The IBIS telescope (Imager on Board of the INTEGRAL Satellite) (Ubertini et al. 2003) onboard the ESA INTEGRAL (INTErnational Gamma-Ray Astrophysics Laboratory) (Winkler et al. 2003), which was launched on 2002 October 17, is a hard-X ray/soft γ-ray telescope based on a coded aperture imaging system (Goldwurm et al. 2001). The coded mask of IBIS is a replicated MURA mask, below which, there are two detector arrays: ISGRI (Lebrun et al. 2003) for low energy band (15 keV – 1 MeV) and PICsIT (Di Cocco et al. 2003) for high energy band (175 keV – 10 MeV).

The size of a detector element is $4.6\,\mathrm{mm} \times 4.6\,\mathrm{mm}$ ($s_d = 4.6\,\mathrm{mm}$), and the mask element size is $11.2\,\mathrm{mm} \times 11.2\,\mathrm{mm}$ ($s_m = 11.2\,\mathrm{mm}$). Thus the angular resolution corresponding to the pixel size of $\mathbf{D}$ is $\arctan(s_d/d) = 4.94'$. Consequently, the intrinsic angular resolution of INTEGRAL/ISGRI is approximately $\arctan(\sqrt{s_m^2 + s_d^2}/d) = 13'$ (Gros et al. 2003).

We use the algorithm introduced in Sections 2 and 3 to some simulated as well as real data of INTEGRAL/ISGRI. In the simulated case, the source image is composed of a flat background and a point source as strong as the Crab, located in the fully coded field of view. In the real data case, the data were retrieved from an observation on the Crab, with the science window 010200210010. The results of the image reconstructions are shown in Figure 2.

This is a typical example showing that DDM can greatly improve the angular resolution, even to much higher than the intrinsic angular resolution of the instrument. The intrinsic angular resolution of INTEGRAL/ISGRI is about $13'$. In contrast, by using the DDM the angular resolution is first improved to about $5'$ (see Fig. 2(c2)) and then to about $2'$ (see Fig. 2(d3)), which is about six times better than the intrinsic one. However, in Figure 2(d3), the intensity of the pixel between the two sources is about 1/10 of the intensity of the sources. This means that these two sources are not completely separated. When the two sources are very close together, it is very hard to distinguish them even by splitting the pixel into more sub-pixels. Thus we infer that the angular resolution of INTEGRAL/ISGRI cannot be much better than $2'$. Namely, the ultimate angular resolution of INTEGRAL/ISGRI is about $2'$.
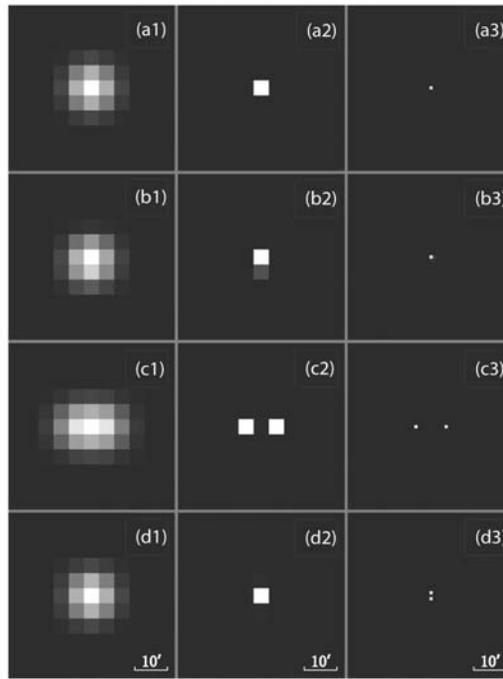
Fig. 2 Images reconstructed by cross-correlation (col. 1), normal procedure of accelerated DDM (col. 2) and improved procedure of accelerated DDM, in which each pixel is divided into 5×5 sub-pixels (col. 3). Images denoted by (a*) are reconstructed from simulation data of one point source. Images denoted by (b*) are reconstructed from real observation of Crab, and (b1) is provided by OSA — the offline scientific analysis software for Integral. Images denoted by (c*) are derived from simulation data of two sources which are $9.88'$ (two detector pixel) away each other. (d*) are images of simulation data of two sources which are $1.98'$ (0.4 detector pixel) away each other.

5 DISCUSSION

Cross-correlation image reconstruction is a linear approach. Its resolution can be simply represented by the FWHM (Full Width at Half Maximum) of a point source. In contrast, DDM is a non-linear method. The FWHM of a point source in the DDM image is much smaller than the smallest distance between two DDM-separable sources. It is therefore reasonable to use the latter as the angular resolution of DDM.

The approach presented in this paper is applicable not only to MURA coded-mask systems like the one onboard INTEGRAL, but also to coded-mask systems with other forms of mask patterns, even patterns to which cross-correlation is not applicable. More generally, it is applicable to shift-invariant systems to obtain images with higher resolution from observational data of lower sampling rate.

It seems that DDM can obtain images which contain more information than the observational data do, but actually this is not true. In the simulation study in Section 4, we only extract the positions and intensities of the sources, plus a smooth background, but not all details of the sky in the field of view. However, it dose extract more information from the data than that by the cross-correlation method, this is why it can separate two $2'$-apart point sources for INTEGRAL/ISGRI data, while the cross-correlation method cannot.

As two sources get close to each other, the difference between their synthetic shadowgram and the shadowgram of one source with their combined intensity becomes smaller. To have distinguishable images, the sources should be so strong that the difference is not submerged in the fluctuations. The sources in the simulations in the former two sections are all strong. At present, due to the limited sensitivity, it is hard to find two sources with separations less than $13'$ in one Science Window of the INTEGRAL data. In the next step, we will further modify the procedure introduced in this paper, so that it will process multi Science

Window data, and both the angular resolution and sensitivity of the images will be improved. We can then reprocess the INTEGRAL Galactic Plane Survey data to find more compact and fainter X-ray sources.

References

Caroli E., Stephen J. B., Di Cocco G., Natalucci L., Spizzichino A., 1987, Space Sci. Rev., 45, 349

Chen Y., Li T. P., Wu M., 1997, proceedings of ADASS VI, ed. G. Hunt, H. E. Payne, ASP Conf. Ser., 125, 178

Chen Y., Li T. P., Wu M., 1998, A&AS, 128, 363

Chen Y., Song L. M., Li T. P., Cui W., 2000, Acta Astron. Sinica, 41, 214

Di Cocco G., Caroli E., Celesti E. et al., 2003, A&A, 411, 189

Feng H., Chen Y., Zhang S. N. et al., 2003, A&A, 402, 1151

Fenimore E. E., Cannon T. M., 1978, Appl. Opt., 17, 337

Goldwurm A., 1995, Exper. Astron., 6, 9

Goldwurm A., David P., Foschini L. et al., 2003, A&A, 411, 223

Goldwurm A., Goldoni P., Gros A. et al., 2001, in Proc. of the 4th INTEGRAL Workshop, ed. A. Gimenez, V. Reglero, C. Winkler, ESA-SP, 459, 497

Gottesman S. R., Fenimore E. E., 1989, Appl. Opt., 28, 4344

Gros A., Goldwurm A., Cadolle-Bel M. et al., 2003, A&A, 411, 179

Lebrun F., Leray J.-P., Lavocat P. et al., 2003, A&A, 411, 141

Li T. P., Wu M., 1993, Ap&SS, 206, 91

Li T. P., Wu M., 1994, Ap&SS, 215, 213

Lu F. J., Li T. P., Sun X. J. et al., 1995, in: Shellard & Nguyen (eds) Proceedings of CHEP'95, 848

Lu F. J., Li T. P., Sun X. J. et al., 1996, A&AS, 115, 395

Lu F. J., Aschenbach B., Song L. M., 2001, A&A, 370, 570

Lucy L. B., 1974, AJ, 79, 745

Paul J., Ballet J., Cantin M. et al., 1991, AdSpR, 11, 289

Richardson W. H., 1972, J.Opt.Soc.Am., 62, 55

Shen Z. J., Zhou J. F., 2007, HEP&NP, 31, 1016

Skinner G. K., 1995, Exper. Astron., 6, 2

Ubertini P., Lebrun F., Di Cocco G. et al., 2003, A&A, 411, 131

Winkler C., Courvoisier T. J.-L., Di Cocco G. et al., 2003, A&A, 411, 1

Zhang S., Li T. P., Wu M. et al., 1997, AIP Conf. Proc., 410, 578

Zhang S., Li T. P., Wu M., 1998, A&A, 340, 62