ELSEVIER

Nuclear Instruments and Methods in Physics Research B 145 (1998) 469–481

NIM B
Beam Interactions
with Materials & Atoms

The EM imaging reconstruction method in γ-ray astronomy

Fernando Jesús Ballesteros Roselló [a,*], Filomeno Sánchez Martínez [a], Víctor Reglero Velasco [b]

[a] Instituto de Física Corpuscular, University of Valencia – CSIC, Dr. Moliner 50, E-46100 Burjassot, Valencia, Spain
[b] Department of Astronomy and Astrophysics, University of Valencia, Dr. Moliner 50, E-46100 Burjassot, Valencia, Spain

Received 11 May 1998; received in revised form 14 July 1998

Abstract

The simpler imaging reconstruction methods used for γ-ray coded mask telescopes are based on correlation methods, very fast and simple-to-use but with limitations in the reconstructed image. To improve these results, other reconstruction methods have been developed, such as the maximum entropy methods or the Iterative Removal Of Sources (IROS). However, such kind of methods are slower and can be impracticable for very complex telescopes.

In this paper we present an alternative image reconstruction method, based on an iterative maximum likelihood algorithm called the EM algorithm, easy to implement and that can be successfully used for not very complex coded mask systems, as is the case of the LEGRI telescope. This is the first time this algorithm has been applied in γ-ray astronomy. © 1998 Elsevier Science B.V. All rights reserved.

1. Introduction

When trying to obtain an image of a celestial source emitting X or γ radiation, classical telescopes based on lenses or mirrors are useless due to the high energy of the radiation [1]; these photons are so energetic that they can pass through the lens without suffering any significant deviation. Grazing-incident reflection [2] is a method that allows to focus the low energy X-rays by striking the photons in multiple reflectant surfaces. This technique implies surfaces whose normal is at great incidence angles with regard to the arriving photons, bigger than the critical angle of the reflecting surface material, deflecting the radiation towards a focus where the detecting surface is. Thus one can form an image in an almost "classical" way. However this technique is more inefficient as the arriving radiation energy increases, because the incidence angle becomes too near to 90°, requiring high amount of surfaces in order to have a reasonable aperture area. As a result, this technique only is feasible for energies below 15 keV.

If we want to obtain images for energies over 15 keV, we need to use a coded mask telescope. It consists of an opaque plate that allows the pass of radiation through a certain pattern of holes placed between the detector plane and the source.

* Corresponding author. Tel.: 34-6-3983262; fax: 34-6-3983261; e-mail: fernando@pollux.uv.es

0168-583X/98/$ – see front matter
PII: S0168-583X(98)00513-8

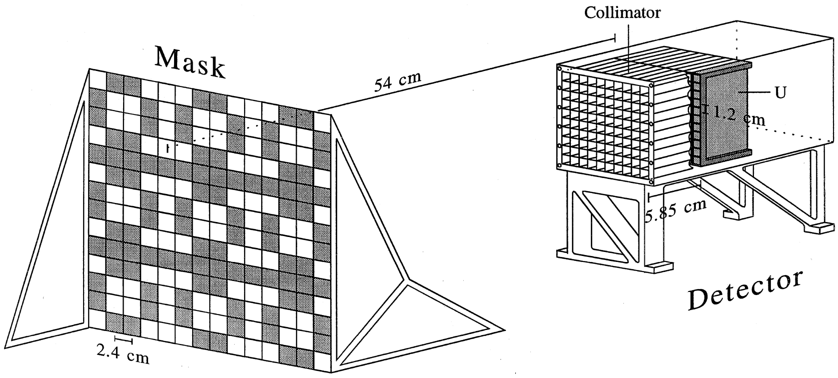

Fig. 1. LEGRI system, showing the detector plane, the collimator and the mask pattern.

Therefore it modulates (codifies) the signal arriving from the source: each γ-ray source casts a shadow of the mask in the detector plane.

Afterwards, original source image can be obtained by deconvolving the recorded image on the detector plane. Nevertheless, the method used to make such deconvolution is very important, as the final quality of scientific data observation strongly depends on the ability of the used algorithm, to properly treat the recorded data.

We have developed an alternative image reconstruction method, based on an iterative maximum likelihood algorithm (EM algorithm) which is used for the first time in γ-ray astronomy. The EM algorithm can be easily implemented for not very complex telescopes, as is the case of the Low Energy Gamma Ray Imager (LEGRI).

The LEGRI is a coded mask telescope optimised for hard X / soft γ-ray domain (20–100 keV) [3] on board the spanish MINISAT-01 satellite. LEGRI consists of a position sensitive γ-ray detector, made up of an array of 10×10 solid state detectors (80 HgI_2 and 20 CdZnTe, ≈1 cm^2 each) with a collimator system in order to limit its field of view, plus a coded mask at 54 cm from the detector plane (Fig. 1). Its continuum sensitivity is 4×10^{-3} ph s^{-1} cm^{-2} in 10^5 s, at a 3σ level. The coded mask pattern is a 5×5 MURA [4] (Fig. 2) placed in a mosaic of 14×14 pixels, each of size 2.4×2.4 cm. The field of view of the instrument due to the collimator system is ±10.5° and the angular resolution is 2.5°.

The main goals of LEGRI are: (i) to demonstrate the technological feasibility of future γ-ray imagers based on solid state detector technology and coded mask techniques, (ii) to perform a galactic survey in order to observe compact objects like black hole candidates, neutron stars, pulsars and high mass X-ray binaries.

We have implemented the EM algorithm for the LEGRI instrument in order to test this method when applied in γ-ray astronomy.

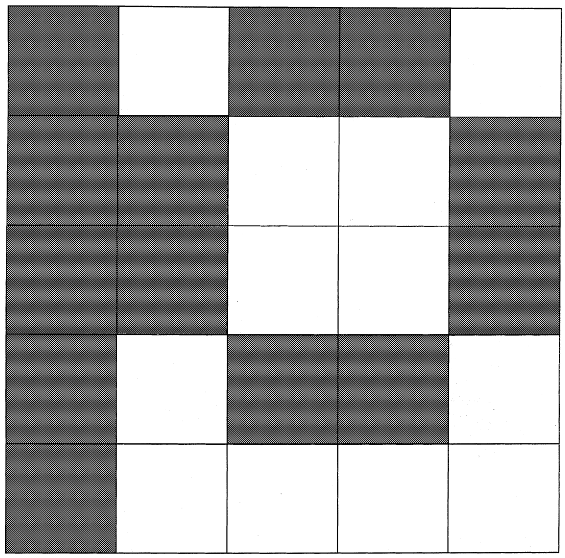

Fig. 2. 5×5 MURA, basic pattern of the LEGRI mask.

2. The EM algorithm

When using coded masks together with pixelled detectors, the recorded image on the detector plane is not a direct image of the source, as in the case of using lenses, but the modulation of the sources by the telescope response, that usually can be approximated by the correlation of the sources with the mask:

$$D_{kl} = \sum_{ij} O_{ij} M_{i+kj+l}. \quad (1)$$

Eq. (1) can be written in a more compact way as

$$D = O * M, \quad (2)$$

where D_{kl} is an array containing the counts detected at the detector kl per unit of time, O_{ij} accounts for the photons emitted from the sky pixel ij per unit of time, and M is an array representing the mask pattern ($M_{ij} = 1$ for a hole and $M_{ij} = 0$ for an opaque element).

To obtain the unknown original image (O) we have thus to process the recorded image (D), by inverting mathematically the response of the telescope. The usual way is by correlating the recorded image with the mask pattern or a modification of it, called the reconstruction array [5]. This method gives good results and is quite fast, and for this reason it is preferred for very complex telescopes. But the accuracy of the method could be fairly bad if the real response of the telescope differs significantly from the theoretical one, given by Eq. (2) (for example, if there are passive structures interfering with the sources, strongbacks, unhomogeneities on the mask transparency, parts of the detector plane that have become unworkable, different efficiencies of the detectors, etc).

To improve the quality of the reconstructed image we have to implement in our model the effect of all the elements involved on the detection process. This can be done if we rewrite Eq. (1) as:

$$D_{kl} = \sum_{ij} \Phi_{kl}^{ij} O_{ij}, \quad (3)$$

where noise term has not been considered (this contribution will be treated later, Section 3.3).

Here Φ is a function that gives the flux (in fact the fraction) detected at the detector kl coming from the sky pixel ij and its value ranges from 0 to 1. In Φ we can include *all* the elements that take part on the radiation detection process as we stated above. Some of these effects cannot be calculated analytically but they can be obtained by Monte Carlo techniques or measuring them directly when possible. Of course, the more realistic Φ is, the better is the reconstruction.

Given this more realistic parametrization of the detection process, one can consider different methods that have been successfully applied in the field of coded mask telescopes for creating a better image, such as the Iterative Removal Of Sources, direct minimization of the χ^2 distribution [6] or the maximum entropy methods [7].

In this work we present an alternative algorithm, which can be successfully applied to reconstruct images when using coded mask telescopes: the EM ("Expected value" and "Maximization") algorithm [8]. It is an iterative by construction algorithm for computing maximum likelihood estimators from incomplete data. This image reconstruction method has been successfully used in nuclear medicine [9], this being the first time it is used in γ-ray astronomy.

The philosophy of the method is the following (see [9]): let us suppose that the data observed in an experiment is a vector D, with an associated conditional probability function $g(D|O)$, where O is a set of unknown parameters to be estimated (the sky pixels in our case); that is, g stands for the probability of obtaining the data D given the parameters O. Our aim is to find the set of parameters $O^{\max}$ that maximizes $g(D|O)$, which will be the best estimator of the real value of the parameters O. In general it is rather difficult to maximize $g(D|O)$ with respect to O; so, instead of working in the observed (and incomplete) data space (which we call $\mathbf{D}$), we will work in the larger space of theoretical complete data, called $\mathbf{D}_\mathrm{t}$, where the optimization will be easier to achieve. The data of this theoretical space, D_t, cannot be directly observed but only through the real data D.

We assume that there is a (non-univocal) mapping $D_t \rightarrow s(D_t)$ from $\mathbf{D}_\mathrm{t}$ to $\mathbf{D}$ and that the D_t values can be known only if they are included in $\mathbf{D}_\mathrm{t}(D)$, i.e. the subset of $\mathbf{D}_\mathrm{t}$ determined by the equation $D = s(D_t)$. We postulate for the complete data D_t also a conditioned probability function $f(D_t|O)$. Under these assumptions it is possible to obtain again $g(D|O)$ from $f(D_t|O)$ by means of the relation:

$$g(D|O) = \int_{D=s(D_t)} f(D_t|O)\mathrm{d}D_t, \tag{4}$$

where the integral is approximated to a discrete sum if when we work with discrete variables (and hence the probability functions are just probabilities).

Each iteration $n+1$ of the EM algorithm consists of two steps; to find out an expected value (**E** step) and to maximize it (**M** step):

- **E** step: to form the conditional expected value

  $$E(\log f(D_t|O)|D, \tilde{O}^n), \tag{5}$$

  where $\tilde{O}^n$ stands for the array of parameters estimated in the iteration n (it has known values, therefore).
- **M** step: to maximize this expected value with respect to O, keeping the values $\tilde{O}^n$ constant. This gives us a new vector of estimated parameters $\tilde{O}^{n+1}$.

The intuitive idea is that we would like to know the parameters O that maximize $\log f(D_t|O)$. Since we do not know $\log f(D_t|O)$, we maximize instead *its expected value* in the present iteration, for a given experimental data D and current estimation of the parameters, $\tilde{O}^n$.

This procedure generates sets of parameters that fulfil the following property [9]:

$$\log g(D|\tilde{O}^{n+1}) \geqslant \log g(D|\tilde{O}^n) \leftrightarrow g(D|\tilde{O}^{n+1}) \geqslant g(D|\tilde{O}^n) \tag{6}$$

being strictly greater in many cases. That is, the EM algorithm is designed to increase the likelihood in each iteration.

Lange and Carson [9] applied this method for imaging in nuclear medicine, for emission and transmission tomography. In the case of emission tomography [9], the D data can be related to the O parameters by means of:

$$D_k = \sum_{ij} \Phi_k^{ij} O_{ij} \tag{7}$$

where k are the different projections of the tomography, ij are the pixels of the source that contribute to the projection k and Φ is an array that represents the probability that a photon leaving pixel ij reaches the projection k (note that in this case the problem is reduced to one dimension, as different tomography projections, D_k, are considered). Due to the *strict concavity* of the function $\log g(D|O)$ the sequence O^n converge to the correct solution $O^{\max}$. The solution for the EM algorithm in this case is given by:

$$\tilde{O}_{ij}^{n+1} = \frac{\tilde{O}_{ij}^n}{\sum_k \Phi_k^{ij}} \sum_k \left(\frac{\Phi_k^{ij} D_k}{\sum_{i'j'} \Phi_k^{i'j'} \tilde{O}_{i'j'}^n} \right). \tag{8}$$

Eq. (8) can be rewritten as:

$$\tilde{O}_{ij}^{n+1} = \tilde{O}_{ij}^n \frac{\sum_{kl} \Phi_{kl}^{ij} \left(\frac{D_{kl}}{\tilde{D}_{kl}^n} \right)}{\sum_{kl} \Phi_{kl}^{ij}} \tag{9}$$

being $\tilde{D}_{kl}^n = \sum_{i'j'} \Phi_{kl}^{i'j'} \tilde{O}_{i'j'}^n$. In Eq. (9) a second index to the detector data D has been added in order to account for the two-dimension detector plane array. As can be seen, it is not necessary to impose additional constraints to ensure the non-negativity of the reconstructed image, except that the initial parameters must be positive, that is, $\tilde{O}^0 \geqslant 0$. This does not preclude the possibility

$$\lim_{n\to\infty} \tilde{O}_{ij}^n = 0. \tag{10}$$

Moreover, it converges to the proper maximum likelihood estimator, independently of the initial value ($\tilde{O}^0$) [9]. In order to avoid introducing any previous structure in the image we will use a uniform field of value 1 for every sky pixel ij as initial parameters.

Eq. (9) is similar to the Lucy–Richardson [10,11] method. However, the Lucy–Richardson algorithm needs to be kept between a lower (0) and upper (O^{MAX}) boundaries in each iteration in order to keep the iterative process under control and to ensure a properly convergency. Nevertheless, the term that arises in the denominator of the EM algorithm assures automatically the convergency of the iterative process without any additional control.

2.1. Convergency and stopping criteria

When an image is generated by means of an iterative process, we have to stop it at a given moment and accept the result of the last iteration as the image generated by the reconstruction process. Therefore it is important to know if the iterative process converges and, if so, when we can consider the process is finished.

The EM algorithm convergency to the maximum likelihood estimator is monotonous and without oscillations as we will show later (Section 3.1); we can use this property to choose a stopping criterion for the algorithm. Nevertheless, this is a criterion about the convergency of the process, not about the convergency to the correct result. Therefore, we can impose other criteria to stop the algorithm, for example, based on the similarity of the estimated data ($\tilde{D}^n_{kl}$) with the real one (D_{kl}). We can stop the process whenever the estimated data (obtained from the present sky intensities estimation) are within the error range of the real data, that is:

$$\tilde{D}^n_{kl} \in (D_{kl} - \sigma_{kl}, D_{kl} + \sigma_{kl}). \tag{11}$$

This is similar to impose conditions on the chi-square value defined as:

$$\chi^2 = \sum_{kl} \frac{\left(D_{kl} - \tilde{D}_{kl}\right)^2}{\sigma^2_{kl}}, \tag{12}$$

where $\sigma_{kl} = \sqrt{D_{kl}}$ (poissonian statistic). Nevertheless, this requirement may be not fulfilled. For these cases we should use the former mentioned monotonous convergency property of the algorithm to stop it. We can stop the iterative process when the algorithm is close to its convergency value, that is, the differences between the images of two consecutive iterations are *very* small. Therefore to account for these differences we define δ^n as:

$$\delta^n = \frac{\sum_{ij} \Delta\tilde{O}^n_{ij}}{\sum_{ij} \tilde{O}^n_{ij}}, \tag{13}$$

where $\Delta\tilde{O}^n_{ij} = |\tilde{O}^n_{ij} - \tilde{O}^{n-1}_{ij}|$. The more similar these two successive images are, the smaller is δ^n. We can impose that the algorithm stops when δ^n is lower than a given value. In practice, we check in each iteration both stopping conditions (Eqs. (11) and (13)), and the process is stopped when one of them is fulfilled.

3. The EM method for LEGRI

In order to test the EM algorithm we have developed a Monte Carlo simulator of the LEGRI telescope, a medium size γ-ray telescope, using the GEANT·3 simulation package [12] together with a simpler and faster geometrical simulator. Full LEGRI geometry and material were considered in the Monte Carlo simulator, taking into account all the physical processes involved in the detection of γ-rays. Different sky sources and background noise levels have been considered in the calculations. In Fig. 3 we show as an example the expected shadowgram obtained in the detector plane when illuminating LEGRI with a centred point source emitting 100 photons cm^{-2} at 100 keV and in absence of background noise. Fig. 3a shows the result for the Monte Carlo simulator and Fig. 3b for the geometrical simulator. The counts detected in the shaded part (where one should expect no counts) in Fig. 3a are due to the scatter of the incident photons with the mask structure. We can see that both simulators produce a very similar result, and therefore we can use the faster geometrical simulator to obtain detector planes to work with.

3.1. Execution time and convergency

Given a pixelled sky map (to be reconstructed by the algorithm) with a size of $A\times B$ pixels, and a detector plane with a size of $a\times b$ detectors (in the case of LEGRI 10×10), the number of operations (see Eq. (9)) for I iterations of the EM algorithm are:

$$[(a \times b) \times (A \times B + 1) + (A \times B) \times (a \times b + 2)] \times I = [2 \times a \times b \times A \times B + 2 \times A \times B + a \times b] \times I.$$

To show the convergency of the EM algorithm, we are going to use the geometrical simulator and illuminate it with a centred point source of 100

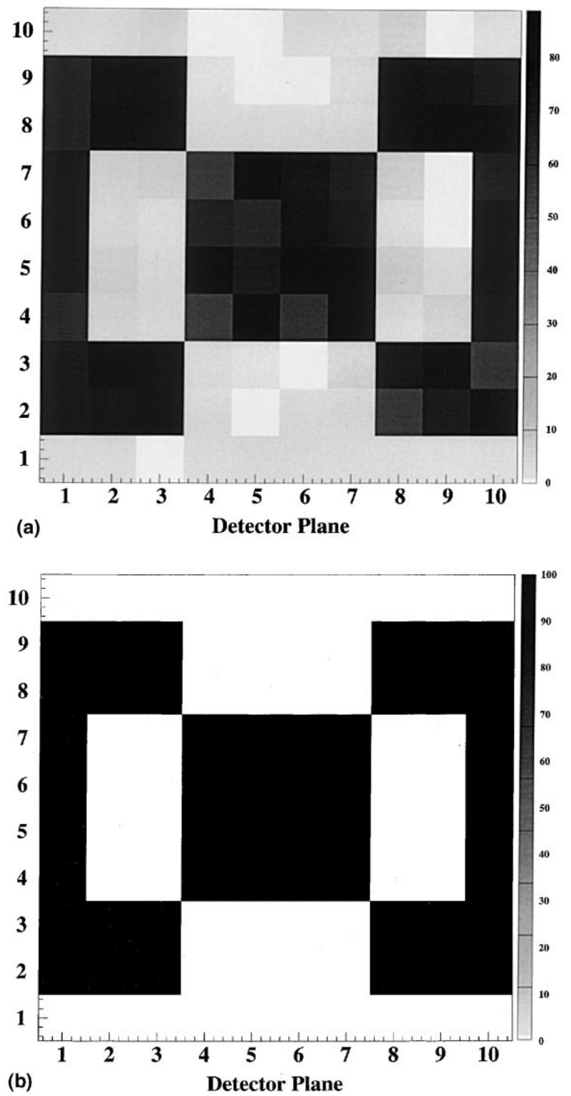

Fig. 3. Expected shadowgrams in the detector plane produced by a centred point source emitting 100 photons cm^{-2} at 100 keV without background noise; (a) using a Monte Carlo simulator and (b) using a geometrical simulator.

photons cm^{-2} plus a random noise with average value of 30 counts cm^{-2}. In Fig. 4 we show the values of δ^n, χ^2 and intensity of the reconstructed source versus iterations, obtained during the EM reconstruction for this example. The algorithm stopped, after 1 min and 25 s (using a Sun Sparc 20 with 120 Mbytes RAM) at iteration 445, fulfilling the stopping criterion of Eq. (11).

The EM algorithm converges very well to the correct solution and does not need any control. Moreover, the convergency of the algorithm is monotonous and achieves differences between consecutive images smaller than 1% after only 50 iterations. The reconstructed image for the centred point source can be seen in Fig. 5.

The number of pixels considered in the reconstructed image of Fig. 5 is $A\times B=$ 33×33 and the detector plane has $a\times b=10\times10$ detector units. This gives 2.2×10^5 operations per iteration.

For the same number of iterations, a more complex γ-ray telescope as is the case of IBIS, one of the main instruments of the future INTEGRAL mission with 128×128 detectors, a field of view of ±14.5° and an angular resolution of 12′ (which gives 580×580 sky pixels if we subdivide it in the same way we have done in LEGRI's image) the EM algorithm will need about 1.1×10^{10} operations per iteration; if we assume the same number of iterations, i.e. 445, the EM algorithm will last about 50 days. This is the main reason why such kind of algorithms can be successfully applied to medium size telescopes but can be unapplicable for very complex γ-ray telescopes.

3.2. Effects of the mask geometry

The mask pattern of LEGRI is a 5×5 MURA pattern [4] placed in a 2.8×2.8 mosaic. The size of each mask element is 2.4×2.4 cm, being the separation between each element of the detector plane of 1.2 cm. Therefore, the detector plane has the same size as the 5×5 MURA. This configuration was chosen in order to be able to use the classical reconstruction methods based on the correlation: any source in the field of view will cast in the detector plane a complete (although permuted) shadow of the basic 5×5 MURA pattern, i.e. it is a cyclic system [13]. In this case it is possible to use a cyclic correlation instead of a simple correlation, and correlate the data with the reconstruction array G as follows:

$$\tilde{O}_{ij} = D \underset{\text{cycl}}{*} G = \sum_{k=0}^{a-1}\sum_{l=0}^{b-1} D_{kl} G_{(k+i)\bmod a (l+j)\bmod b} \qquad (14)$$

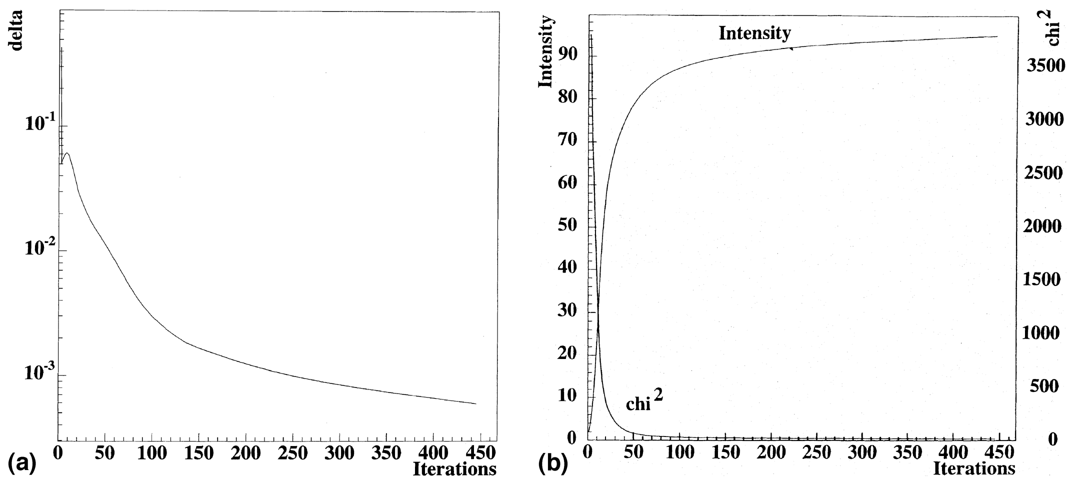

Fig. 4. Convergency of the EM method for the magnitudes (a) δ^n, and (b) source intensity and χ^2.

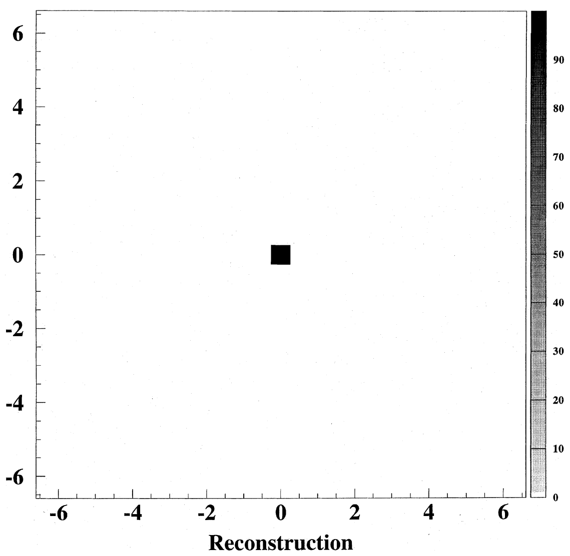

Fig. 5. Reconstruction of a 100 ph cm^{-2} centred point source.

being a and b the size of the detector plane and the reconstruction array. Using Eq. (2) it is equal to:

$$\tilde{O} = D * G = O * M * G \quad (15)$$

where $M * G$ should be a delta function in the ideal case. This will be possible if the system is a cyclic one and the basic mask pattern is a URA [1] or MURA pattern, as in our case.

This kind of straightforward methods are very fast and for this reason are preferred for very complex telescopes. But for medium size telescopes, we have seen that the EM algorithm is competitive in computing time.

3.2.1. Angular resolution

In the classical reconstruction methods based on the correlation, the angular resolution is given by the angle subtended by a mask element viewed from the detector plane, that is:

$$\text{resol.} = \text{arctg}\left(\frac{c}{f}\right) \quad (16)$$

being c the size of a mask element and f the distance mask–detector plane (for LEGRI, $c = 2.4$ cm and $f = 54$ cm, obtaining an angular resolution of 2.54°). This resolution, given by the correlative methods, is usually the nominal resolution of the coded mask-based γ-ray telescopes and it is independent of the spatial resolution of the detector plane.

Obviously, if the detector plane has a good spatial resolution (i.e. can resolve pixels 10 times smaller than the mask elements), it is expected in principle to have a better angular resolution. Therefore, one has to conclude that when using a correlative method, there is implicitly a loss of information. If we use the EM method, we can retrieve this lost information and obtain the real angular resolution:

$$\text{resol.} = \text{arctg}\left(\frac{d}{f}\right) \quad (17)$$

being d the size of the detector plane pixels. The real *angular* resolution on the telescope depends on the *spatial* resolution of the detector plane. For LEGRI, this value is 1.27°. Only in the case when $d = c$ there is no loss of information when using a correlative method.

We can show it easily if we apply the field of sources shown as crosses in Fig. 6 to our LEGRI simulator (without any background in this case) and we reconstruct the detected image by both, a correlative method and the EM algorithm. In this case we are limiting ourselves to a central zone of the field of view (±6°).

In Fig. 6a we have reconstructed the image using a correlation-based method called δ-decoding [14] which is a correlative method with a contrast greater than other correlative methods. In Fig. 6b, we have used the EM algorithm. The image obtained with the EM algorithm has better contrast and resolves better the different sources (except those two too near in the centre, because their angular separation is lower than 1.27°).

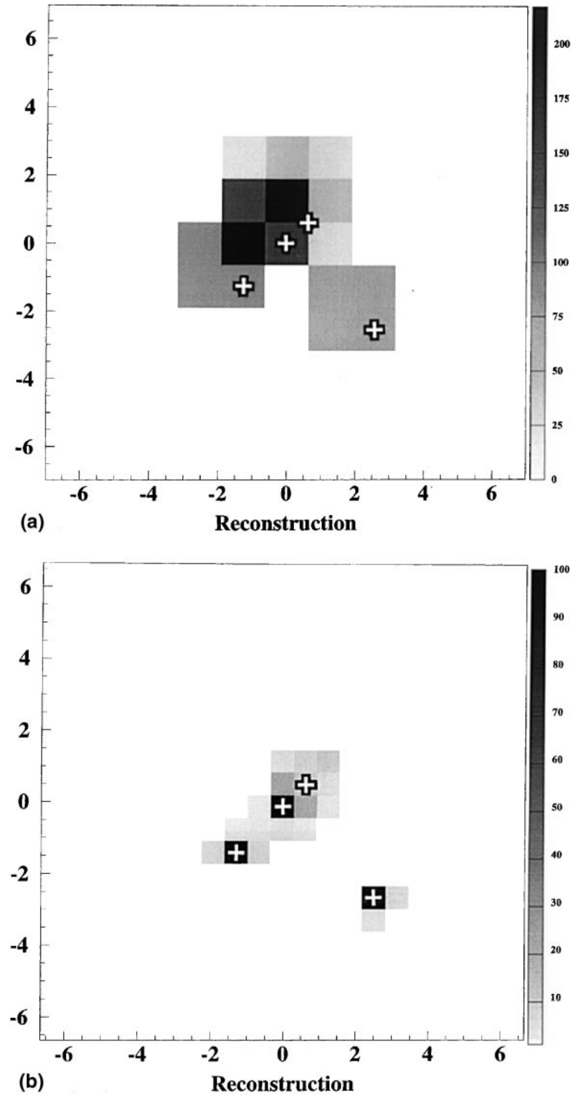

Fig. 6. (a) Reconstruction using δ-decoding (a correlative method); (b) reconstruction using the EM algorithm. Shown only the central part of the image. In both cases the sky sources used as input in the simulator are marked with crosses.

3.2.2. Ghosts

Using a cyclic system such as the one of the LEGRI telescope, with the basic pattern in a 2.8×2.8 mosaic, one has the disadvantage that *different* source positions on the sky can cast *the same* pattern in the detector plane. This will give an intrinsic degeneration in the system and one will be unable to distinguish between certain source positions. As the reconstruction methods will not have additional information to distinguish among those directions, this fact will produce fake sources in the reconstructed image that we call *ghosts*.

To show up this effect, we have applied the sky source distribution marked with crosses in Fig. 7 to our simulator. After applying the EM algorithm to the detected signal, we obtain the reconstructed image shown in Fig. 7. We can see that the central source has been reconstructed without any ghost, because the shadow cast by a source on-axis cannot be produced by any other source position (due to the collimators surrounding each detector units which limit the field of view to ±10.5°). But this is not the case for the other sources. The source located at (5,0) produces two sources in the reconstructed image, one in the real position and the other source (ghost) located just in a position where it would cast the same shadowgram as the real one. The difference in reconstructed intensities is due to the triangular response of the collimators.

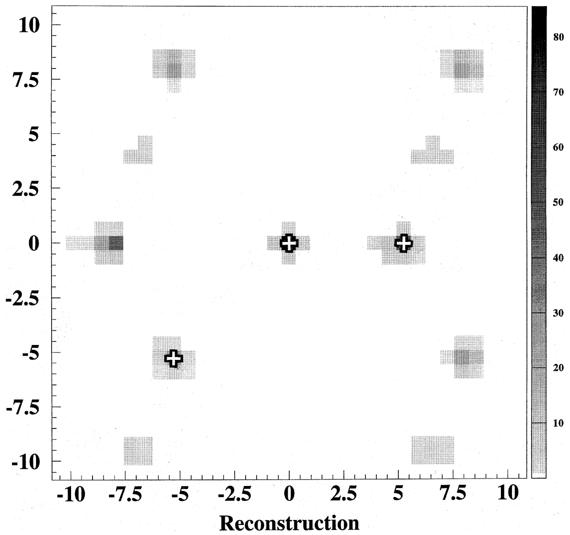

Fig. 7. Reconstruction using the EM algorithm, showing the real input sources (marked with crosses) and their ghosts.

The same situation occurs for the source located at (−5, −5), which produces three ghosts, corresponding to the other possible positions at which a source would produce the same shadow on the detector plane as the true one.

This effect of "ghosting" can be avoided by using a non-cyclical mask (although in this case it is not possible to use a cyclic correlation), for example a random mask or a bigger URA or MURA masks.

3.3. Noise effect

In general, added to the signal coming from the sky source, one has a background noise that affects to the signal. There are different sources of noise that can be considered, i.e. other sky sources, the diffuse γ-ray background, cosmic rays, protons and electrons trapped in the earth's magnetic field (in the case of LEGRI's L.E.O. orbit, the south Atlantic anomaly -S.A.A.-), radioactivity induced by these energetic particles in the structure of the telescope, electronic noise, etc... all of them are just counts added to the γ-ray sky source counts we want to study.

The easiest way to introduce the background noise in our calculations is just to add it to Eq. (3):

$$D_{kl} = \sum_{ij} \Phi_{kl}^{ij} O_{ij} + B. \tag{18}$$

In Eq. (18) we have considered that background is independent of the considered detector. In a more realistic situation, each detector will detect a different background, not only due to differences in the efficiency, but also due to the fact that, as the telescope structure is activated, the detectors closer to this structure will have higher background counting rate. In this case we could have a situation similar to that shown in Fig. 8 where, added to a central point source of 100 photons cm^{-2}, we have included an anisotropic background, bigger in the border (∼300 counts cm^{-2}) where the strongback structure of the telescope is, and decreasing towards the centre (∼50 counts cm^{-2}).

Taking the possibility of anisotropic background, we can rewrite Eq. (18) as:

$$D_{kl} = \sum_{ij} \Phi_{kl}^{ij} O_{ij} + B_{kl}. \tag{19}$$

If we assume we know the background noise (we have a good model of it, B^{model}), the usual way to treat with it when we are looking at a sky γ-ray

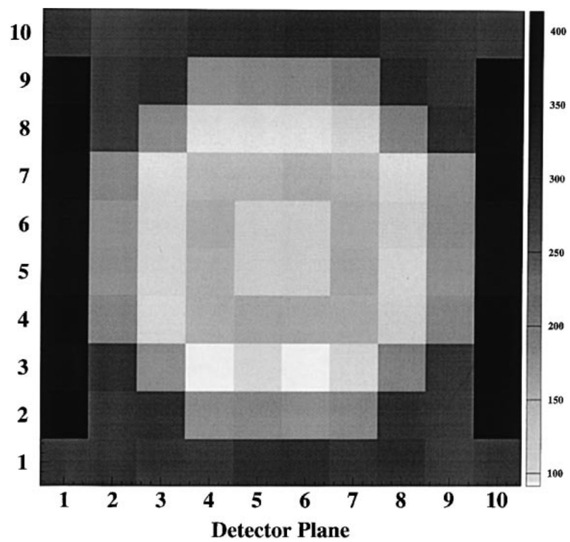

Fig. 8. Shadowgram in the detector plane produced by a 10 ph cm^{-2} centred point source plus an anisotropic background.

source is just to subtract it from the data, in order to have an equation similar to Eq. (3):

$$D_{kl}^{\text{corrected}} = D_{kl} - B_{kl}^{\text{model}} = \sum_{ij} \Phi_{kl}^{ij} O_{ij}. \tag{20}$$

Then, if we use correlative methods, we can apply our reconstruction array directly to $D^{\text{corrected}}$. But unfortunately we cannot act in the same way if we use the EM algorithm, because the subtraction $D_{kl} - B_{kl}^{\text{model}}$ can produce negative values in $D^{\text{corrected}}$ (there are detectors measuring only background due to the mask shadow, and statistical fluctuations implies that the measured counts can be lower than the estimation of the background, B_{kl}^{model}). To assure positive values of the estimated source intensities $\tilde{O}$ (Eq. (9)), D_{kl} must be positive.

The correct way to treat with the background when using the EM algorithm is to consider it *as a part of the detector response*. Therefore, the estimators of the D values will include implicitly a background term:

$$\tilde{D}_{kl} = \sum_{ij} \Phi_{kl}^{ij} \tilde{O}_{ij} + B_{kl}^{\text{model}}, \tag{21}$$

the way to obtain this model of the detector plane background, to include it as part of the detector response, is to look at a field of view without any γ-source (and so the sum for ij in Eq. (21) is equal to 0) and measure the detected counts.

If we use as input Fig. 8 (non-homogeneous background plus a source emitting 100 ph cm^{-2}), the results in the reconstruction can be seen in Fig. 9. In Fig. 9a we are not considering any model of the background noise; in Fig. 9b, we subtract before reconstruction a model of the background noise (a model similar to the background we have introduced previously) to the detected counts, according to Eq. (20); in Fig. 9c, we add the model of the background as a part of the detector response, according to Eq. (21). Sidelobes arise when background is not considered (Fig. 9a). In Fig. 9b the EM algorithm produces an even worse result with negative values for some pixels. Finally, in Fig. 9c the sources are properly reproduced when background is considered as a part of the detector response (Eq. (21)).

3.4. Non-perfect detector plane

When we reconstruct an image by means of a correlative method, we have to correlate the data with the so-called reconstruction array G according to

$$\tilde{O} = D * G = O * M * G + B * G \tag{22}$$

or what is the same

$$\tilde{O}_{ij} = \sum_{kl} D_{kl} G_{k+i\ l+j} = \sum_{i'j'} \sum_{kl} O_{i'j'} M_{i'+kj'+l} G_{k+i\ l+j} + \sum_{kl} B_{kl} G_{k+i\ l+j}, \tag{23}$$

where G should satisfy that $M * G$ is a delta function (the identity for the correlation operation) and $B * G$ as close as possible to 0. In this equation is implicitly the fact that all the D_{kl} data have the same importance and that we have to take into account all of them in order to have a correct image reconstruction: if some detectors are switched off or damaged – and so $D_{kl} = 0$ for those detectors – , then some of the values of M that should be different from 0 will be 0 and therefore $M * G$ will not be a delta function anymore. To test the effect on the reconstructed image of this non-perfect detector plane, we have used as input of the simulator the same field of sources of Fig. 6. We have switched off three rows of detectors and eight more detectors selected at random, that is 38% of the detector plane has been disabled. The detector plane shadowgram can be seen in Fig. 10.

When we reconstruct the image using a correlative method, we get the image of Fig. 11a where it is not possible to distinguish any source. On the other hand, if we apply the EM algorithm to this seriously damaged detector plane we obtain the result showed in Fig. 11b. The result obtained by the EM algorithm is rather better although some odd structures arise in the reconstructed image.

However we can still improve this result: we should not consider the information coming from them as valid information (as we do if we just apply the EM algorithm to the detected counts) because we know they do not work. The way to avoid this incorrect treatment of information from

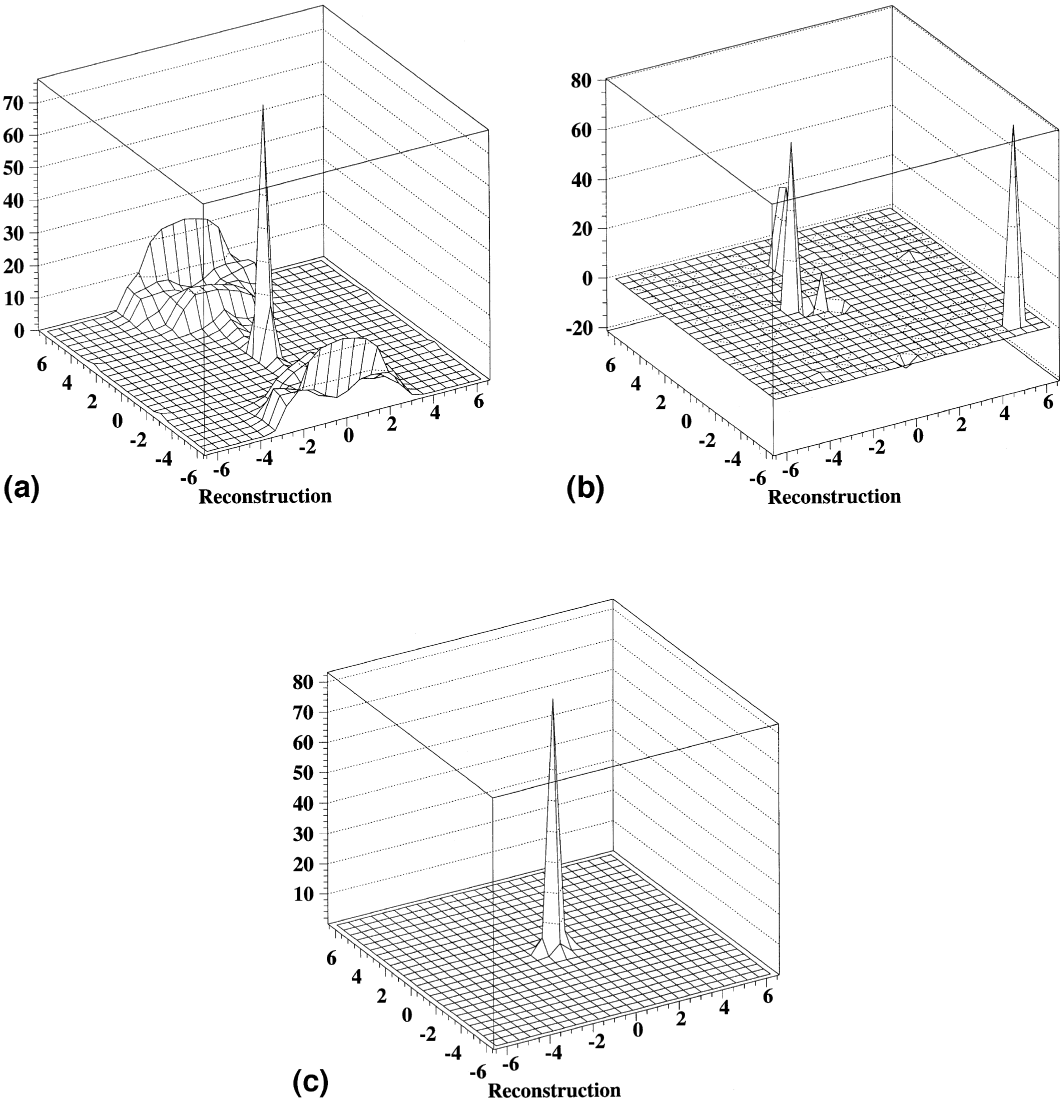

Fig. 9. EM reconstructions of Fig. 8: (a) without considering any model of background; (b) subtracting a model before reconstruction; (c) adding the same model as a part of the detector response.

these detectors is by means of a by-pass of the disabled or damaged detectors. That is, in Eqs. (9) and (21) we have to exclude from the sum all the damaged detectors:

$$\sum_{kl} \rightarrow \sum_{kl\,\notin\,\text{damaged detectors}}. \qquad (24)$$

In this case we obtain the result showed in Fig. 11c, which is better than the previous one and very similar to the result we obtained when we had the whole detector plane working (see Fig. 6c). That is, we can recover more information from the source with the EM algorithm than that obtained with the classical correlative methods (notice that this *by-pass* is done *automatically* in the case we use a correlative method, so we cannot improve the correlation results by means of this by-pass).

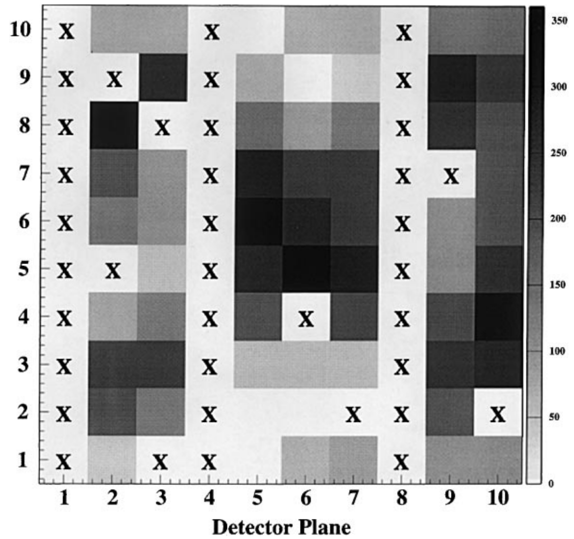

Fig. 10. An example of damaged detector plane, illuminated by the same sky sources of the case shown in Fig. 6.

4. LEGRI data and preliminary results

When we work with the real data coming from the LEGRI instrument we have to deal with a very important background noise, which makes the data handling more difficult. This noise is due to the fact that Minisat 01 is in a Low Earth Orbit (L.E.O.) with a height of 550 km which passes across the South Atlantic anomaly. This fact induces great amounts of activation in the LEGRI structure material, producing a strong background which dominates over the sky data.

On the other hand, we have a severe damage due to problems in the launching and the effect of the strong radiation environment where LEGRI is orbiting, and about the 80% of the detector plane has become useless. As a result of these two effects, the sensitivity of LEGRI is not good enough to see the main part of the sky gamma sources.

Anyway, with the measurements we have made nowadays, we can see the brightest sources in the sky, as the case of the pulsar in the Crab Nebula. Using the data we have obtained from the Crab Nebula and applying them our EM algorithm, we obtain the image reconstruction shown in Fig. 12. We can see at the centre the signal from the pulsar in the Crab Nebula, and surrounding it, six ghosts

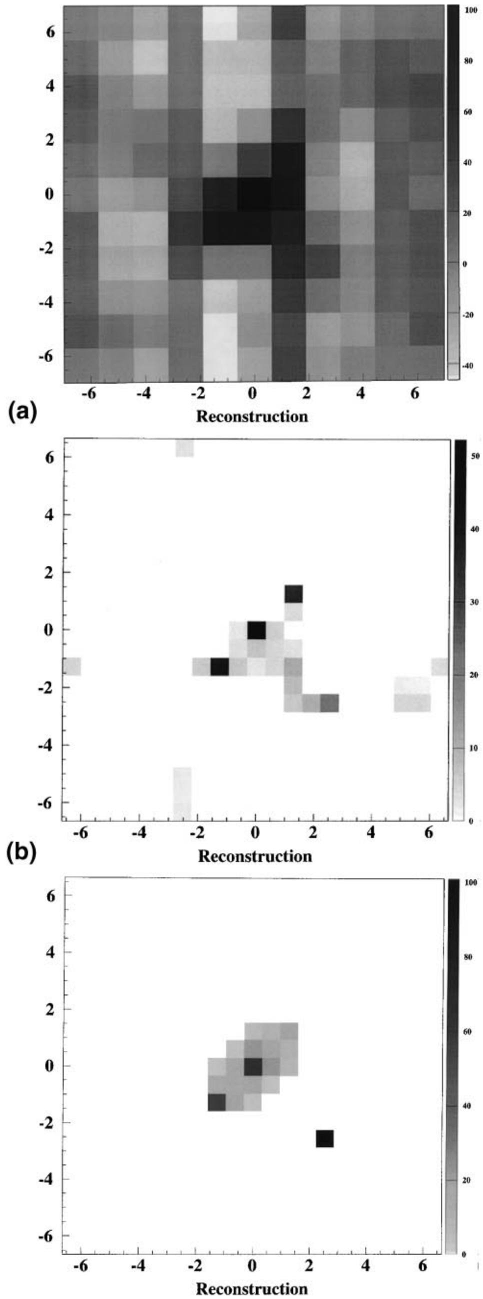

Fig. 11. Reconstructions of Fig. 10: (a) using a correlative method; (b) using the EM algorithm; (c) using the EM algorithm and by-passing the damaged detectors.

or fake sources due to the mask pattern and uncertainties produced for the damaged detector plane. Due to the strong limitations in pointing

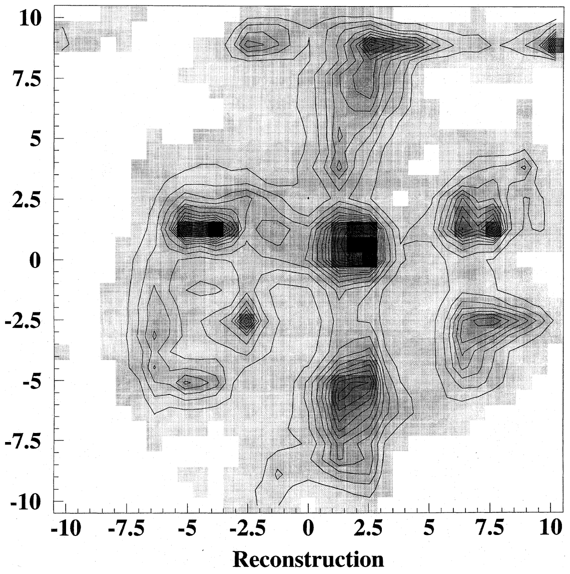

Fig. 12. Reconstruction of the Crab Nebula pulsar from LEGRI real data.

and observing time of the platform Minisat 01, we cannot increase the integration time as quickly as we would like, in order to improve the sensitivity of the instrument, and only acquiring more data in the future we will increase the integration time and so the sensitivity of the instrument.

5. Conclusions

This is the first time the EM algorithm has been applied to coded-mask based γ-ray telescopes.

The EM algorithm is a powerful iterative method for image reconstruction, with very good convergency characteristics: robustness, no need of additional controls or checks to assure convergency, positivity of the solution and convergence to the proper solution. The EM algorithm has a suitable computing time for medium-size coded-mask based telescopes which makes it competitive with the correlative methods. This algorithm has been applied to LEGRI telescope, on board Minisat-01, showing a very good adaptability to difficult conditions (i.e. space-varying background noise in the detector plane, improvement of the angular resolution, very high background noise rate in the real data, an important damage in the detector plane), giving as a preliminary result an image of the Crab Nebula pulsar. It is expected that future measurements will increase the integration time of the sources and therefore the sensitivity, allowing to detect weaker sources.

With the EM algorithm it is possible to recover more information from the data than that obtained with classical correlative methods. The EM algorithm seems to be a very suitable method for image reconstruction of medium size γ-ray telescopes where coded mask techniques are needed.

References

[1] G.K. Skinner, Scientific American, August 1988.
[2] L.P. Van Speybroeck, R.C. Chase, Appl. Opt. 11 (1972) 440.
[3] V. Reglero et al., in: Proceedings of the second INTEGRAL international workshop "The Transparent Universe", 16–20 September 1996, St. Malo, France, p. 343.
[4] S.R. Gottesman, E.E. Fenimore, Appl. Opt. 28 (1989) 4344.
[5] E.E. Fenimore, T.M. Cannon, Appl. Opt. 17 (1978) 337.
[6] G. Ducros, R. Ducros, Nucl. Instr. and Meth. 221 (1984) 49.
[7] T.J. Ponman, Nucl. Instr. and Meth. 221 (1984) 72.
[8] A.P. Dempster et al., J.R. Statist. Soc. B 39 (1977) 1.
[9] D. Lange, R. Carson, J. Comput. Assist. Tomogr. 8 (2) (1984) 306.
[10] L.P. Lucy, Ast. J. 79 (1974) 745.
[11] J. Richardson, J. Opt. Soc. Am. 62 (1972) 55.
[12] GEANT-Detector Description and Simulation Tool, Application Software Group, Computing and Network Division, CERN, Geneve, Switzerland, 1994.
[13] G.K. Skinner, Nucl. Instr. and Meth. 221 (1984) 33.
[14] E.E. Fenimore, T.M. Cannon, Appl. Opt. 20 (1981) 1858.