# Combining GAIA windows I: SNR calculations of secondary sources in the ideal case for AF11

P. Nurmi*

DMS-PN-01, October 29, 2003

*Royal Observatory of Belgium, Av. Circulaire 3, 1180 Bruxelles, Belgium*

**Abstract**

During 5 year observational phase of GAIA, several one-dimensional observations are obtained from the same objects at different epochs and transit angles. It is scientifically important to combine all the observations to produce a two-dimensional image that contains all the data. This allows us to find much fainter stars surrounding the main source than what would be possible to find in single observations on-board the satellite or in individual transmitted patches. Most of the emerging objects are physical secondary stars, but also some background/foreground objects can be found. The combination procedure also increases the image resolution in the cross-scan direction. The procedure consists of combining patches that are stretched back to 2-D images when the original centroid positions of windows are known. Transit angles, i.e. image rotations, can have values between 0° and 360°, which are generally not uniformly distributed. The initial image window from which patches are taken is not a square, but it is elongated along scan-direction in the Astro AF11 field of view, that is used in this study. The window combining method is a simple mapping procedure that is fast and can easily handle all realistic complexities occurring during GAIA data reduction. The method is not intended to be a perfect image restoration technique, but the purpose is to fulfill the scientific requirement of detecting faint neighboring sources existing in the close vicinity of the main point sources. For this purpose a simple mapping procedure is suitable. In this study we show examples of patch combination and illustrate the basic reduction procedure that could be used to detect faint companions around primary targets. We aim searching the ultimate limits that can be achieved using the method and discuss possible complexities occurring during the final data analysis. Using GIBIS 1.4 images as an input we estimate the efficiency of the method and conclude that at the distances of a few arcseconds the method can reveal very faint objects down to G=24 and is especially efficient. But close to a bright companion Poisson noise starts to dominate and source detection becomes very difficult and a more elaborate model fitting techniques should be used.

*E-mail: Pasi.Nurmi@oma.be

# 1 GAIA imaging analysis

2D imaging of the surrounding of a detected star can be obtained by means of the 1D observations from the whole 5 year mission. An 'observation' from one CCD of GAIA consists of a window containing one or several patches. A 'patch' is a 1D set along scan of equally large samples. A 'window' consists of one or several patches (see figure 1). We treat in this report only observations from the CCD AF11 in the astrometric field, and since the window for AF11 consists of only one patch, the words patch and window are fairly interchangeable.

The basic requirement of the patch combination procedure is that signal-to-noise ratio (SNR) can be increased and hence faint stars and other sources, otherwise hidden to noise, can be detected. Also, the effects of varying pixel size can be removed and the resolution of the images can be increased.

Scientifically, detection of faint background stars or binaries having larger magnitude differences is very important issue. During the final photometry reduction all disturbing sources around the main target must be known to achieve the best possible accuracy for positions and magnitudes. Stacking of all the patches at different transits and combining a super-image that contains all the information from all the transits, makes it possible to detect sources that are hidden to noise in single images. At moderate separations (1-10 pixels) for double stars, we can detect binaries having large magnitude differences (up to $\Delta m = 8$). Outside of the dirty PSF (term is defined in section 2.2), close to the edges of the windows, $\Delta m$ is limited only by the magnitude of secondary sources. If the magnitude and the position of the secondary object are the main properties we are interested in, it is best to use unfiltered images from Astro. The Astro field AF11 provides a large field of view (unfiltered patches over 2.5") in which also relatively distant neighboring stars can be detected. However, only overlapping pixels provide additional data and therefore the best area for window stacking and magnitude gain is determined by the window size across scan-direction.

# 2 Window combining procedure

## 2.1 Basic idea

The method used for window combining is loosely based on the variable-pixel linear reconstruction or drizzling, that is a procedure in which input pixels are sub-sampled to some output grid that have smaller $1 \times 1$ pixels (by fraction $f_1$) than in the initial frame (Fruchter and Hook 2002). Before actual mapping, initial pixels are shrank (by fraction $f_2$) or some other procedure (e.g. Gaussian weighting) is used so that the final image is not smoothed too much (Gilliland et al. 1999). Since the drizzling method was initially designed for combining undersampled images, pixel shrinking was needed, but for GAIA images the PSF is well sampled and we keep the final pixel scale the same as for samples along scan-direction. For AF11 windows the pixel scale used in the combined image is therefore 44.2 mas. In the following analysis we use pixels as a unit of length, but pixel values can always be converted to arcseconds by multiplying

pixel values by the pixel scale. The initial idea of Fruchter and Hook is also modified so that it can handle elongated 1-D pixels i.e. 1-D patch is stretched back to original size of the window before mapping of samples is done.

To describe the method we first fix the notation used in the following description. The notation agrees with the general GAIA terminology (Bastian 2003). Initially a 2-D window $IM_{t,k,m}$ at an epoch $t$ is obtained at the satellite in AF11, where $k$ is window pixel coordinate along scan-direction and $m$ is pixel coordinate across scan-direction. Due to data transfer limitations only part of the initial pixel data in the CCD is transmitted to ground and therefore a window sized $\Delta K \times \Delta M$ pixels is generated around the object on-board the satellite.

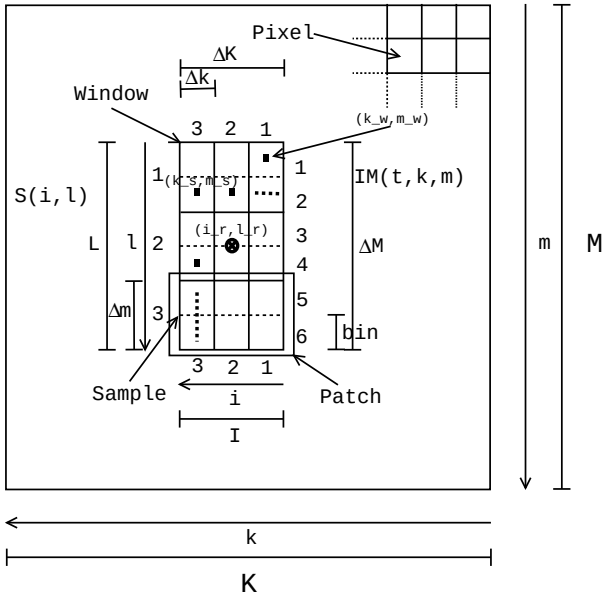


Figure 1: *A general description of the Gaia terminology used in the text.*

Inside the window samples are generated by summing pixels in various ways depending on the target classification criteria. In the window situated at $(k_w, m_w)$ a sample $S_{t,i,l}$ at time $t$ is

$$S_{t,i,l} = \sum_{k=k_s-\Delta k}^{k_s} \sum_{m=m_s-\Delta m}^{m_s} IM_{t,k,m}, \tag{1}$$

where $\Delta m$ is the number of pixels summed together across scan-direction before read-out noise (RON) is included. $k_s$ and $m_s$ are pixel coordinates of the samples in the $IM$. Both, $\Delta k$ and $\Delta m$ depend on the magnitude G and the instrument module. For AF11, $\Delta m = 14$ and $\Delta k = 1$. It should be noted that $\Delta m$ is also used to mean interval of magnitude. In addition to on-chip binning it is also possible to bin samples numerically. In AF11 the numerical bin width varies according to the distance from the center of the main source. When numerical binning is applied, the signal is divided equally between binned samples.

In general $S_{i,l}$ is new sampled image, where $i$ and $l$ are sample coordinates of the samples in the window. We define that $I$ is the total number of samples in the patch and there are $L$ patches in the window. The window size in the units of samples is therefore $I \times L$. For AF11 $I$ is 66 or 68, depending on the primary G-magnitude and $L = 1$. A detailed description of windows and patches in different instrument modules is given in technical report CUO-117 (Høg et al. 2003). A general description of the terminology is shown in the figure 1.

Window combination is done in four basic steps. At first patches are stretched to correspond to the initial size of the 2-D window, from which the patches were taken. Next, the center of the window is shifted in $i$-direction and $l$-direction so that absolute positions of the windows are perfectly coinciding with their positions in the underlying map. The central point of the map (COR), is a fixed point in the sky. In the real case all stars are moving in the final map, but in the ideal static case the primary star is always at the center of the map. After centering windows, the window is rotated by $\theta$, the angle that is obtained from the scanning law and some fixed direction in the sky.

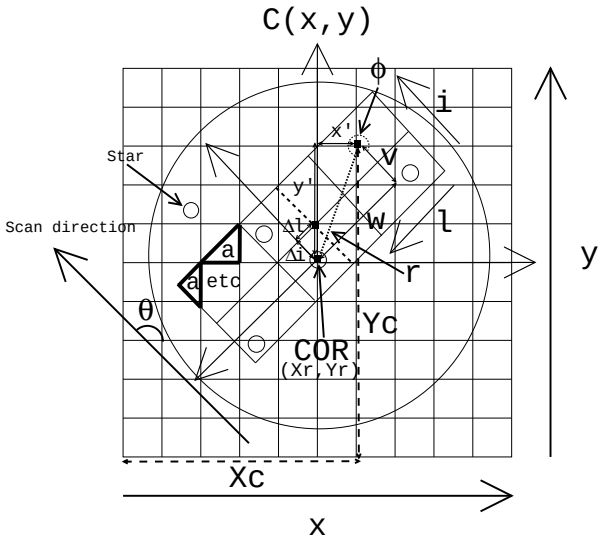


Figure 2: *A general description of the window combining terminology used in the text.*

Since the centroid is generally located close to the center of the window we fix the window center to

$$i_r = I/2 \tag{2}$$

$$l_r = (L/2)bin, \tag{3}$$

where $bin$ is the real size of the sample $bin = 3 \times \Delta m$. A rotation by angle $\theta$ and translation by $\Delta i$ and $\Delta l$ give a new position ($x_c$ and $y_c$) of the pixel center in the final image $C_{xy}$:

$$v = (i - 0.5) - (i_r + \Delta i) \tag{4}$$

$$w = (l - 0.5) \times bin - (l_r + \Delta l) \tag{5}$$

$$r = \sqrt{v^2 + w^2} \tag{6}$$

$$\phi = atan2(w, v) \tag{7}$$
$$x' = r\cos(\phi + \theta) \tag{8}$$
$$y' = r\sin(\phi + \theta) \tag{9}$$
$$x_c = x_r + x' \tag{10}$$
$$y_c = y_r + y', \tag{11}$$

where $i$ goes from 1 to $I$ and $l$ goes from 1 to $L$. Final combined image is $C_{xy}$, where $x$ and $y$ are pixel coordinates. The reference system may be defined in such a way, that x-direction and y-direction are parallel to some fixed orientations in the sky. The scan angle $\theta$ is the angle between the x-direction and scan-direction. Definitions of all the variables are given in the figures 1 and 2.

The values for $\theta$, $\Delta i$ and $\Delta l$ are available when the scanning law of the satellite is known and when source position in the window is known. Pixel positions are shifted so that the window is always on the new image. In the calculations we have used the final image size of 100 pixels $\times$ 100 pixels, that is chosen to be large enough to allow windows to stay within the image in all cases.

After the initial step of positioning the centers of samples, rotated and stretched patches must be combined using some image combination method or more elaborate image restoration techniques. One possibility is to use an inversion technique as described in IC-CDAB-01 by Dollet and Bijaoui. In this study we use the method, in which individual scans are combined using a pixel mapping procedure. Obvious benefits of this method is that it is relatively easy to include different sampling schemes, correction factors and other complexities existing in the real GAIA data reduction. Also, it is possible to evaluate noise levels and signal gain in the final combined image.

The kernel of the algorithm, that is repeated until all the patches are combined, is

$$W_{xy}^{k+1} = a_{xy,x_cy_c} w_{x_cy_c} + W_{xy}^{k} \tag{12}$$
$$C_{xy}^{k+1} = \frac{a_{xy,x_cy_c} i_{x_cy_c} w_{x_cy_c} s^2 + C_{xy} W_{xy}^{k}}{W_{xy}^{k+1}}, \tag{13}$$

where $C_{xy}^{k}$ is combined intensity at $k_{th}$ iteration in the final image, $W_{xy}^{k}$ is combined weight and $w_{x_cy_c}$ is individual weight of the pixel, $a_{xy,x_cy_c}$ is geometrical weight (overlapping factor), $s$ is the fraction between the size of the output pixel to the input pixel ($s = 1/bin$) and finally $i_{x_cy_c}$ is the measured pixel value at $(x_c, y_c)$. If needed, different weights $w_{x_cy_c}$ can be introduced to individual pixels so that geometrical distortions or any relevant corrections are taken into account. In the equations (11) and (12) $x$ and $y$ refer to pixel positions in the output grid. When $a_{xy}$ is calculated the elongated shape of the pixel and partial overlaps with the pixel and the underlying $C_{xy}$ image frame must be taken into account. These calculations are fairly simple and are not explicitly given here.

Equally important for the final image reduction and additional source detection is to understand the noise behavior in combined images. Following Fruchter and Hook 2002, the variance of the noise in the combined image assuming that

the signal is equally distributed over the initial pixel *bin* is

$$\sigma^2_{C_{xy}} = \frac{\sum_{i_{xy}\in C_s}(a_{xy,x_cy_c}w_{x_cy_c}s^2)^2\sigma^2_{x_cy_c}}{(\sum_{i_{xy}\in C_s} a_{xy,x_cy_c}w_{x_cy_c})^2}, \tag{14}$$

where $\sigma_{x_cy_c}$ is the standard deviation of the noise distribution of the input pixel $i_{x_cy_c}$, which overlaps with a set $C_s \in C$. Sub-pixel centers are calculated after the elongated pixel is subdivided into *bin* sub-pixels. Hence, in general for k different noise sources $\sigma^2_{x_cy_c} = \sum \sigma^2_k$. Total noise in the pixel is therefore

$$\sigma^2_{x_cy_c} = \sigma^2_S + \sigma^2_B + \sigma^2_D + \sigma^2_R, \tag{15}$$

where $\sigma_S$, $\sigma_B$, $\sigma_D$ are Poisson noises $\sqrt{e^-}$, where $e^-$ is the number of electrons per sample from the source, background and dark current, respectively. Read-out noise is $\sigma_R = R$. In the noise calculation, the number of electrons from the main source is obtained from GAIA PSFs and background estimation is done for each field separately. Using previous noise calculation (equations 13 and 14), we can calculate the "noisemap" i.e. $\sigma_{C_{xy}}$, that gives the expected noise variations over the combined image. The signal-to-noise ratio (SNR) in each pixel position is

$$SNR_{xy} = \frac{S_{xy}}{N_{xy}} = \frac{n^{e^-}_{xy}}{\sigma_{C_{xy}}}. \tag{16}$$

In this equation we have not included the photon noise from the secondary source itself and therefore for bright secondary sources the calculated SNR values are too high. In the combined image, noise is clumpy due to correlation of noise in adjacent pixels and this is important if the SNR is calculated over many pixels. The problem is that real noise for pixels that are summed together is not $\sum \sigma^2_{x_cy_c}$ due to the correlation of noise. The standard noise calculation gives too optimistic values and SNR values are too high if the signal is integrated over a large area. Therefore, as a first approximation before we have full understanding about noise correlation effects, we calculate SNR as an average value from 4 pixels, although the signal in this area is less than 10% of the total signal. The signal value used in the SNR calculations is therefore very close to the peak-to-background level value measured at the centroid position.

## 2.2 Combined image

As an example, the combined image after two AF11 patch windows are combined is shown in the figure 3. Two arrows indicate two different scan directions and two boxes are drawn around windows from which patches are obtained. In the figure 3 two different co-centered circles are shown. The size of the large window is 68 samples wide in scan direction having 14 combined pixels in one sample. For faint stars there are 66 samples in the patch with 14 combined pixels per sample. Solid line circles are outer limits for the combined image e.g. for large window $r = \sqrt{(3\times 14/2)^2 + (68/2)^2} \sim 40$ pix and $r = 39$ for small window. Dashed-line circle denote the area inside which the pixels overlap completely in every scan and the increase in SNR is at maximum. The size of the

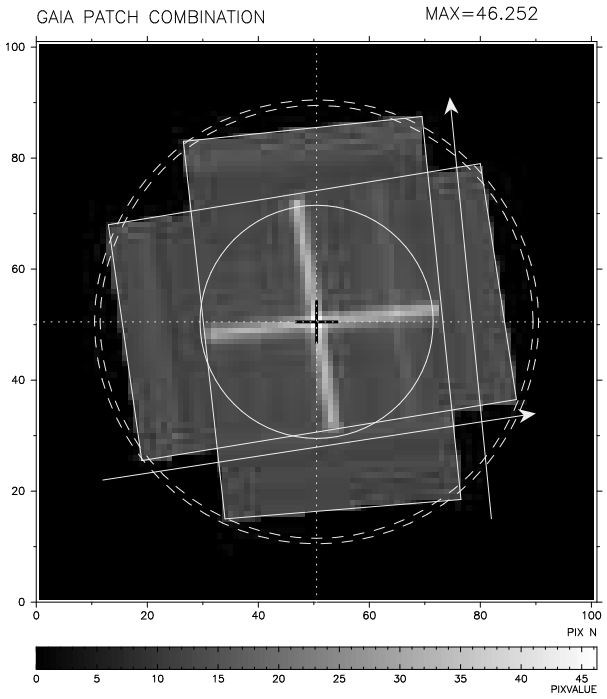


Figure 3: *The main features of the image obtained by the window combining method.*

overlapping region is determined by the window size in across scan-direction and therefore $r = 21$ pixels.

We can also create an equivalent to the PSF in the combined image that we call "dirty PSF". It is created using LSFs, or PSFs in windows having more than one patch per window, that are combined using same shift-values and rotation angles as for observed images. Two examples are given in the figure 4, where the dirty PSF is situated at the center of the image and in another panel the central position is shifted 15 pixels away from the center. In both cases the number of combined windows is 80 and the scan angles are random. The shape of the dirty PSF changes as a function of the number of transits, position in the window, scan angle distribution and the position in the focal plane. In addition, there are internal variations due to magnitude and the color of the star. So the dirty PSF is different for each object and must be calculated for every target separately.

## 2.3 Different phases of image processing for faint source detection

We summarize here the main steps of the proposed method to search for faint sources surrounding the main target.

1. Map stretched patches to underlying grid and combine all patches using the mapping algorithm

2. Obtain LSFs of the central hypothetical single star using all the information that is available and affects LSF (i.e. positions, magnitude, color

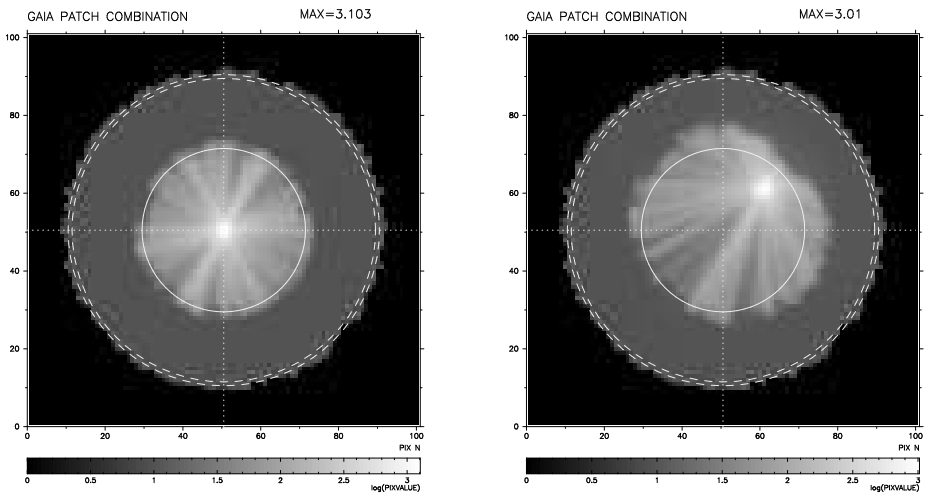


Figure 4: *Two dirty PSFs are shown in the image at different positions in the combined image. 80 windows are combined and the rotation angles are random.*

etc.)

3. Combine LSFs using the same algorithm and COR shifts as used for the observed patches

4. Scale combined images and subtract the dirty PSF image from the main frame

5. Apply a source detection algorithm to final subtracted image using the noisemap for noise estimation to search for faint sources surrounding the primary target

The last four steps are required to detect anything with large $\Delta m$ around the central component and they are all normally performed, since we generally do not know what is hidden in the images. In the ideal combination case, when all pixels are completely overlapping, the flux increase $\propto \sqrt{N}$, where N is the number of combined images. The magnitude gain would then be

$$\Delta m_{gain} = 2.5 \log \sqrt{N}. \tag{17}$$

In the ideal case, this means a difference of 1 mag in 6 images, 1.25 for 10, 1.85 for 30 and 2.22 for 60 images. However, in reality, this is not exactly true and especially for faint objects the signal in the images is readout noise limited and the proportional factor is $< 2.5$, but the gain is still proportional to $\sqrt{N}$. Systematic errors like, calibration errors, can also lower the gain factor, but in any case there should be a notable increase in SNR in the combined image.

Conceptually the method applies to any combination of windows (ASM, AF, BBP, MBP) and it needs to be studied what windows should be combined to maximize the scientific output. It is obvious that the best starting point for faint source detection is to use AF11 windows, since spatial coverage and photon collecting power is the best for AF11. The next step, if something is detected in

AF11, would be to combine BBP patches and search for color difference between the main source and auxiliary sources. In BBP windows there are some factors that may cause the auxiliary sources to remain undetected. Across scan size is not as large as in AF11 and for very faint sources situated away from the central region, the signal might be too low to have reliable detection.

### 2.4 Image subtraction using dirty PSF

The main purpose of image subtraction using dirty PSF is to see faint sources around bright objects. To do subtraction properly between theoretical combined image and the observed one, high SNR is required and good knowledge about the center and the wings of the LSF is required. These requirements are normally fulfilled, since before the window combination we already have fairly precise information about the main target. It is assumed that calibrated LSFs of the main target in the window corresponding to each observation epoch and position in the focal plane are available.

For a small number of transits the binning can be a severe problem in the detection capability, since the magnitude gain can be small outside the central area, but normally $\sim 50$ transits is enough to have a remarkable gain in SNR. If LSF changes as a function of position on the chip and the spectrum of the target or if there is any change in the LSF shape in time (focus etc.) that is not known a priori, then subtraction with "wrong" dirty PSF can create some spurious errors and false detections. A detailed analysis of the residuals in the subtracted image could therefore be used as a confirmation that stellar properties are coherent with the calculated values.

## 3 Test runs

A test software that reads AF11 patches derived from GIBIS 1.4 (version 19.5.2003) (Babusiaux et al. 2001) images as an input and a program that combines the windows of all the transits using the centroid position and transit angle information, has been developed and first test runs for ideal case (perfect psf fitting, no cosmics, random transits) has been done.

### 3.1 Main assumptions for ideal case study

The main assumptions and idealizations made during the window combination procedure and source detection are:

1. LSF is well defined from the color of the star and does not vary in time
2. Centroid position of the primary star is known at sub-pixel accuracy
3. Centroid position of the primary star is random within the pixel in both directions for each transit independently
4. All astrometric motion is ignored
5. Dirty PSF subtraction of the primary star is perfect

6. Source detection algorithm, after main component is subtracted, is perfect

Noise sources and error estimates included in the simulations are:

1. Background and dark current values as in GIBIS 1.4
2. Poisson noise as in GIBIS 1.4 images
3. Read-out noise $R = 6e^-$ is included after proper re-binning is done to images
4. Expected along scan-direction position errors per scan of the stellar centroid position in AF11 are included
5. Since the across scan-direction position error per scan is not easily estimated the error value that is 10 times larger than in along scan-direction is adopted
6. Expected flux errors per scan of the central star for AF11 are also included

Position and magnitude errors are LSF fitting errors in one scan as given by Frédéric Arenou (personal communication).

Other noise sources, ignored in this study, but to be added to the simulations in later stage that can change the detection limits to higher values are:

1. Dirty PSF subtraction errors
2. Astrometric motion
3. Influence of cosmic rays
4. Complicated systems
5. Detection errors and false detections

### 3.2 Description of test simulations

Using GIBIS 1.4 we have generated binaries having at maximum of 100 transits using random transit angle distribution and random sub-pixel positioning. Primary magnitudes of V=12, 16, 18 and 20 with V-I=0 are used in the simulations (for V-I=0 $G \approx V$). For secondary stars we have used magnitude differences of 2, 4, 6 and 8 magnitudes. Mutual separations are between 1 and 7 pixels and in addition, outside the complete overlapping pixel area, we have two test cases at 15 and 30 pixel separations from the main component. The total number of systems is therefore $4 \times 4 \times 9 = 144$ systems and totally this does $1.44 \times 10^4$ transits. The analysis is done using 5 different numbers of transits: 5, 10, 50, 80 and 100 to have an idea how SNR improves as a function of number of transits.

In reality, the number of transits varies greatly according to target' s position in the sky. A quick look to real GAIA scan number-distribution is given in the figure 5. Two different stellar distributions are used: in the first case we have chosen 200 stars randomly from the sky and in the second distribution 200 brightest stars in the sky are used as an input distribution. The latter

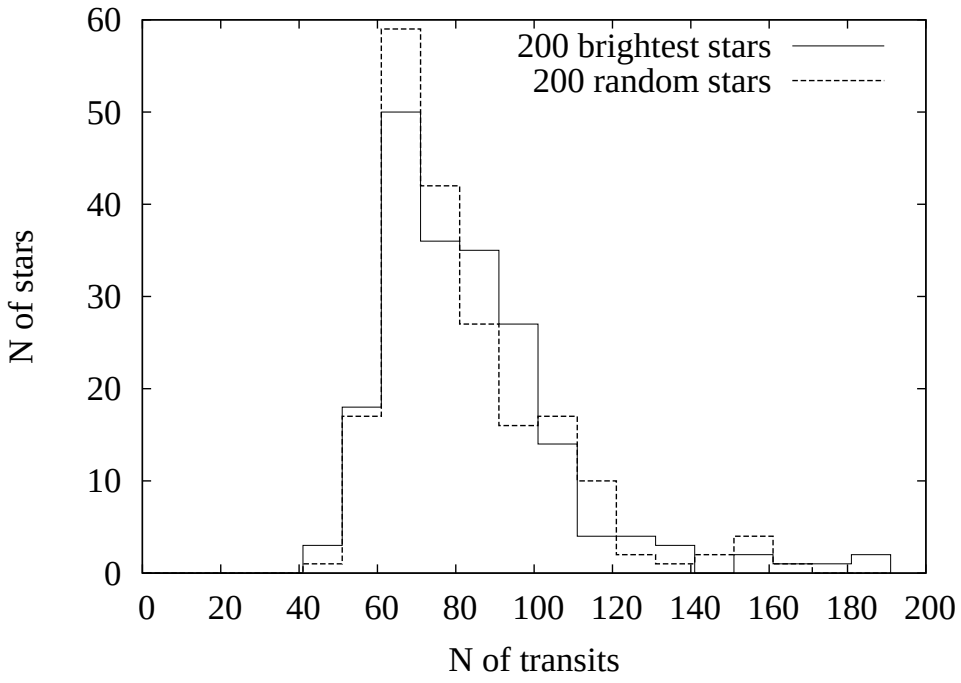


Figure 5: *Frequency histograms for two different initial stellar sky distributions: for 200 randomly positioned stars and the 200 brightest stars in a real-sky catalogue.*

gives more realistic average stellar distribution. However, it turns out that the distributions differ only slightly and for the random case mean and median values of the number of transits are 78 and 83, considerably. Similar values for 200 brightest stars are 75.5 and 82. Based on these values we estimate that the average number of transits is $\sim 80$. All the calculations are done using prevailing GAIA-2 design parameter values.

In this report we are interested in the ideal case and therefore we used only random scan angle distributions, although generally transit angles are not randomly distributed. Some examples of realistic transit angle distributions are given in the figure 6 to six different $l$ and $b$ values. A few points about realistic scan angle distributions should be mentioned. Generally we have more than 50 transits, that is normally enough to achieve a remarkable increase in SNR. Normally when the number of rotations $< 100$, the scan angle distribution is fairly uniform. Some low level peaks are always present, but they should not cause a big problem. Only when the number of transits is large $> 100$, there are two distinct peaks in the distribution (5th panel), separated by $180^\circ$ or then there is only one clear peak (1st panel).

To show the effects of having one or two dominant angles in the scan-angle distributions two subtracted images are shown in the figure 7. Scan-angle distributions used in the images are the same as in the figure 6 (1st and 5th panel). The separation is 7 pixels and the system is $G = 18$ and $\Delta m = 4$. The images are for 100 combined windows from which dirty PSF is subtracted. The overall effect of having one or two dominant scan angles is that in the images

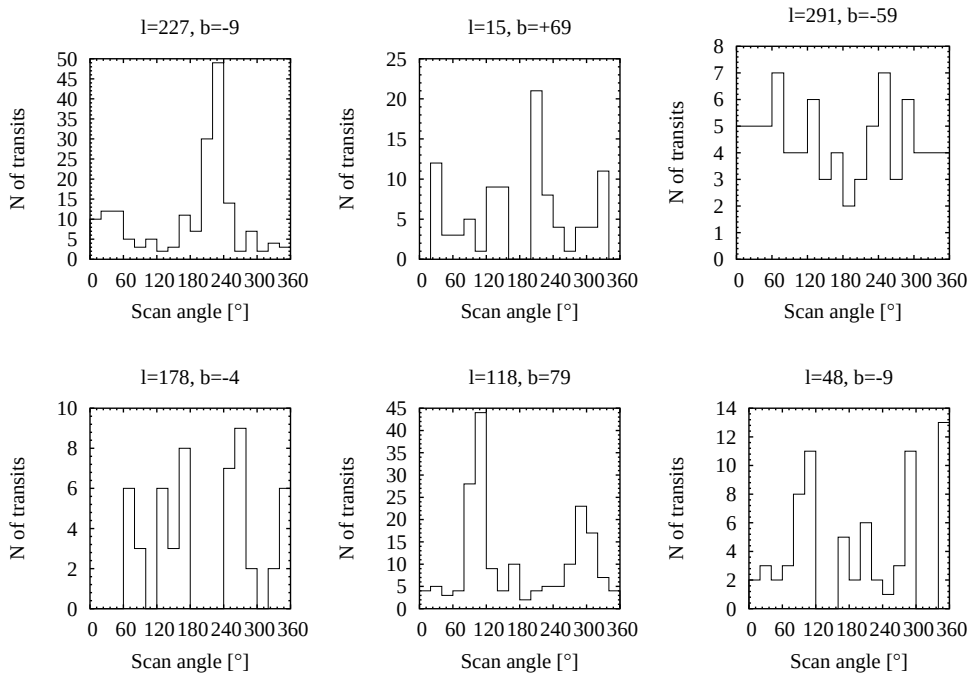


Figure 6: *Six realistic transit angle histograms are shown as calculated using the GAIA scanning. The total number of transits in the six cases from upper left to lower right panels are: 181, 95, 81, 52, 188 and 72.*

there are some dominant noise patterns visible and point sources are elongated in across-scan direction corresponding dominant angles. Generally there is still a clear peak value, but the total signal is spread over a wide area and the magnitude estimation is more difficult.

# 4 Simulation results

## 4.1 SNR behavior for single stars

At first, it is interesting to consider how the SNR increases for faint stars in the ideal case, when no disturbing sources are nearby. This helps us to understand the "empirical" SNR behavior for very faint objects and the limits obtained from this study determine the absolute limits that can be achieved when other sources are present. Therefore, we study how the SNR changes as a function of the number of combined images N, for 5 different magnitudes G=[21-25] and compare these to theoretical expectations. In practice we calculate the signal as an average value from 4 central pixels, from which background has been subtracted. The noise value is taken from theoretical noise expectation values given by the noisemap (eq. 13). Empirical SNR curves are given in figure 8 and they seem to follow nicely theoretical expectations of $\sqrt{N}$ law.

Both, theoretical and empirical SNR curves show that SNR increases rapidly up to $\sim 50$ images, after which an increase in SNR is much smaller. Since the average number of transits for Astro-1 and Astro-2 together is 80 for GAIA, we

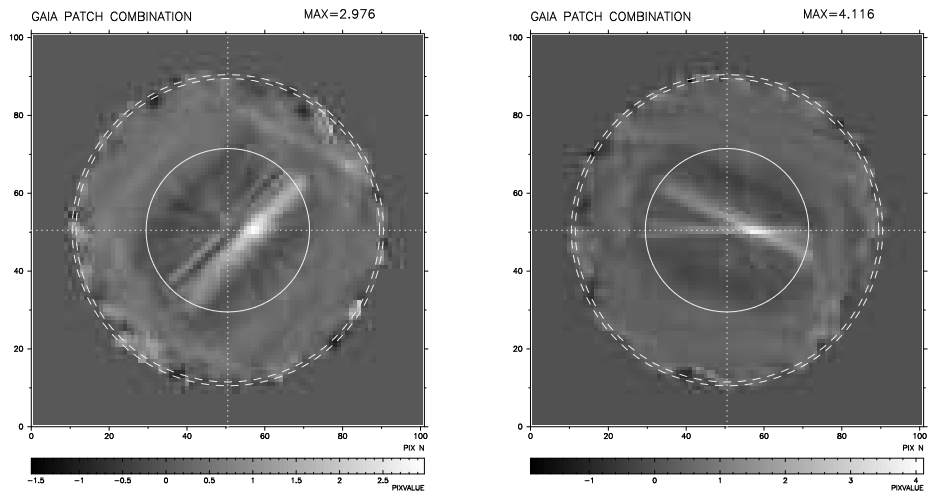


Figure 7: *Two subtracted images of a system with* $G = 18$ *and* $\Delta m = 4$ *using two different realistic Gaia scan angle distributions are shown. In the left panel the scan angle distribution has one dominant scan angle and in the right panel there are two dominating angles having* $\sim 180°$ *difference.*

conclude that generally we have sufficient number of images to increase SNR considerably. On the other hand, it may be that it is not necessary combine all AF11 transits, since the benefits are not that large compared with the increase of calculation time, if it turns out to be a problem in the end. We conclude that at 24th magnitude the detection SNR limit is still at acceptable level (SNR=5-6) and the number of false detection should be small. Hence, the absolute limit that we can achieve by image stacking using prevailing GAIA noise estimations is 23 or 24, depending on the SNR limit that is chosen.

### 4.2 SNR calculations of secondary sources

Faint secondary source detection procedure includes two main steps. At first, the bright component is subtracted away from the image. In the final data handling software this step requires a procedure similar to PSF fitting and to do this we need to know the color and the initial magnitude of the star as precisely as possible. Next step after image subtraction is to study the residuals for possible additional sources. If separation is so close that no a priori information about duplicity is available, then our results are too optimistic. In this case the background source detection depends primarily on the quality of dirty PSF subtraction.

Calculated SNR curves of secondary sources after the main source is subtracted away are given in the figures 9 and 10 for four different primary star magnitudes and four different magnitude differences as a function of mutual separation. Shapes of the SNR curves depend mainly on two competing factors. Close to the main target Poisson noise in high increasing the total noise, but inside the overlapping pixel region, the increase in signal is the highest.

At close separations, for bright objects (primary G=[12,16]), Poisson noise is much higher than signal increase due to combination of patches and SNR

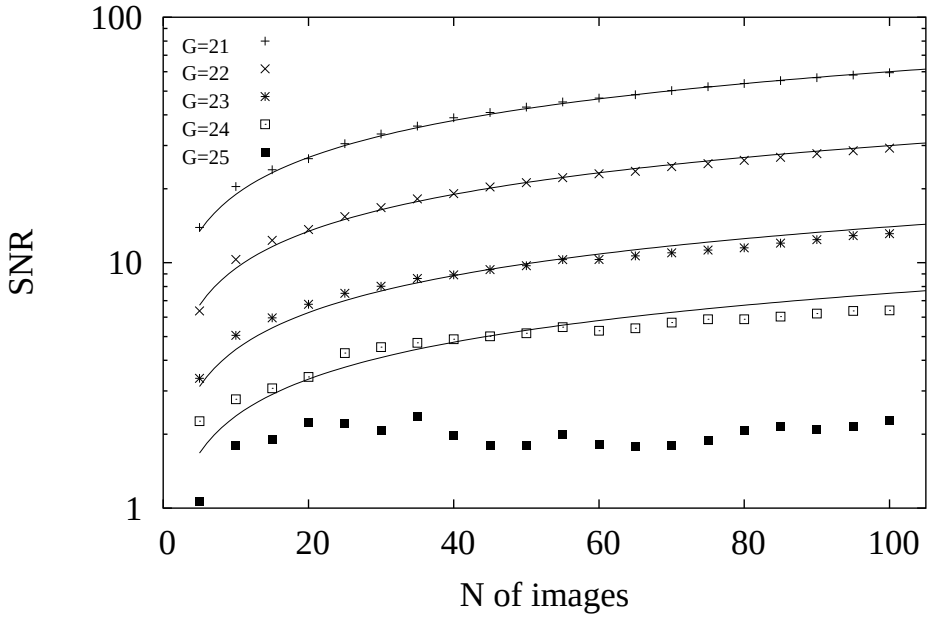


Figure 8: *Calculated SNR curves for five different single stars situated at the center of the window. These curves represent ultimate limits that can be achieved using the combination procedure for single stars in the area where pixels are completely overlapping.*

increases as the secondary moves away from the main source. SNR curves for G=12 and G=16 are given in the figure 9, for five different number of transits 5, 10, 50, 80 and 100.

For faint primaries G=[18,20] in the figure 10, it can be seen that when secondary is situated a few pixels away from the main source, read-out noise dominates the noise in the image and SNR curve has a maximum value at some separation from the primary. The maximum occurs at the point where the object has maximum number of overlapping samples and the Poisson noise is the smallest.

Based on the two previous figures we can set the preliminary resolution and detection upper limits in the ideal case. We conclude that for G=12 it is possible to detect all background sources for all separations and magnitude differences that were studied down to $\Delta m = 8$. For G=16 at $\Delta m = 8$ we can detect only sources that are far away from the main source $\theta > 15$ pix. As was noted for single stars, the ultimate magnitude limit for source detection in the central region of the image is G=24 so at G=18 $\Delta m = 8$ is far too faint to be detected, but when $\Delta m = 6$ we may have some low SNR detections. For G=20, when $3 < \theta < 15$ pix, it is still possible to detect objects with $\Delta m = 4$. High SNR values at $\theta = 30$ when $\Delta m > 2$ are only due to a statistical fluctuation, since the error in SNR is very large.

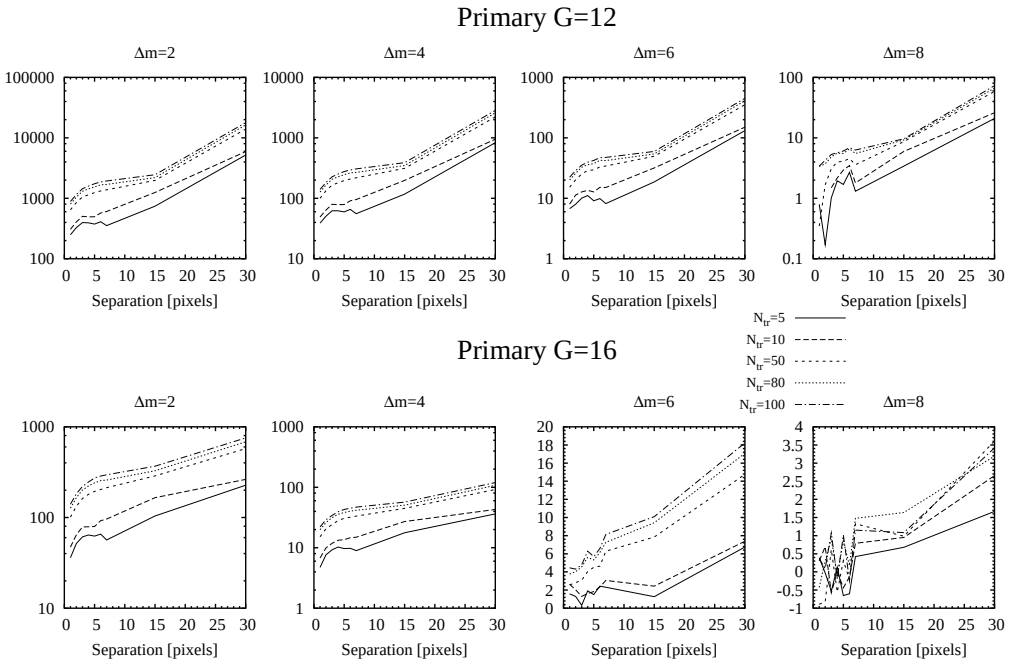


Figure 9: *Calculated SNR curves of secondary sources as a function of separation from the main source are given for two (G=12 and G=16) GAIA primary magnitudes. Five different curves show the SNR variation as a function of number of transits in the ideal (random) case.*

### 4.3 Artefacts and further complications

It may be that the initial color of the bright object was given a wrong value. The effect of this to the source detection is that subtraction using wrong LSFs induces ring-like structures around the source, since the shape of the wings depends closely on the color of the source. In reality, we believe that this should not be a big problem since the color is normally known accurately enough. Further study of this issue is needed in later stage.

Other realistic complications including astrometric motion between the main object and other sources should be also considered. Parallax values are generally sub-pixel level, but for some high proper motion stars and physical binaries with $P < 10$ yr signal from secondary sources can be spread over many pixels and therefore they may remain undetected. Finally, calibration errors and all systematic effects that are not modeled in this study may have an effect on the final analysis. The overall influence of these factors is very difficult to estimate at this point of data analysis design.

We have also run a few simulations using realistic scan angle distributions. These calculations show that for a typical case when the number of transits is $\sim 50$ and the scan angles are not centered around some single value, SNR values does not differ notably from the random case. In the following studies this point also requires more detailed analysis.

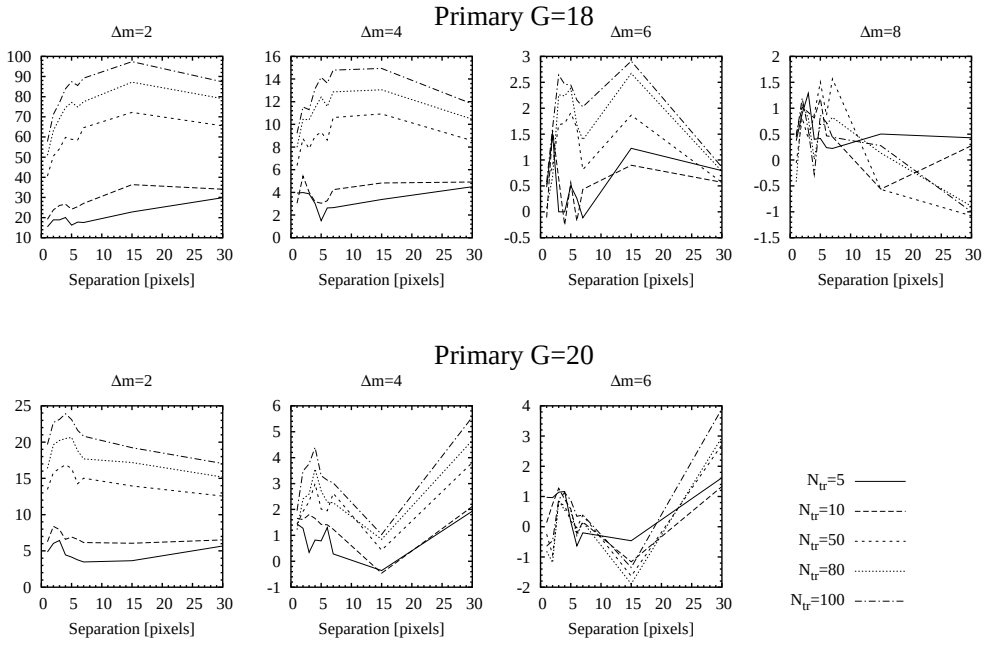


Figure 10: *Calculated SNR curves of secondary sources as a function separation from the main source are given for two (G=18 and G=20) GAIA primary magnitudes. Five different curves show SNR variation as a function of number of transits in the ideal (random) case.*

### 4.4 Requirement of processing power

The processing speed of the combining procedure is far from being optimized and the current speed for patch combining is $\sim 25$ patches/s using 1Ghz desktop computer. This means that the total processing speed is $2 \times (80/25)$s =6.4s for a median case with 80 transits. There is a factor of two since also theoretical LSFs are combined to form a dirty PSF for each target. As a starting point this is acceptable, since we know that the final speed can be greatly improved.

During GAIA data analysis the number of windows to be processed is very high. If all the stars are to be analyzed, then the total number of patches to be handled is of the order of $10^{11}$ and the procedure must be as fast as possible. Even spending one second per target, would mean data reduction time of 32 years. Hence, still a suitable data reduction time during the final data reduction should be $\sim 0.01$s per target, so that the combination procedure does not slow down the overall data reduction performance. If we can now do the analysis in six seconds per object then the final processing speed should be increased 600 times. If we believe that the processing speed increases as predicted by Moore's law during the next 9 years, then the processing speed of desktop computers in 2012 would be 64 times higher than presently. This means that the combining procedure of all targets could be performed within a few months using a small PC cluster available in 2012 even using the code currently available.

## 5 Discussion

What information is then obtained from the combined images? The output depends a lot on the capabilities of the final data reduction procedure and probably different categories of source detections depending on the SNR and the type of object should be described. The final image reduction algorithm should be able to provide at least secondary positions and magnitude estimations including errors of the parameters. For some objects an estimation of the color from BBP images could be obtained. This is limited to much smaller field of view and brighter stars than in AF11. For very faint and close-by secondaries we can probably only state that the object is "complicated" from the residuals, but more detailed analysis is not possible. On the other hand in "clear" cases and for objects that we have a priori knowledge to expect that there should be something additional in the image, much more detailed analysis can be done using probably also MBP patches.

For visual binaries with small orbital periods the number of overlapping pixels is much smaller than for stationary objects and therefore faint close-by companions could be missed. The motion of the secondary (if detected at all) should form trails to combined images. If we have a priori knowledge about the secondary, a different algorithm that takes the motion into account could be applied and a detection could be possible. Fortunately these objects should be detectable by astrometric fitting procedure and hence some additional information could be at hand. Also photometric binaries and spectrophotometric binaries that have very close separations are more easily detected by the methods created specifically to these objects.

Lastly, I would like to point out some connections to other tasks that I see to be relevant to my work. In the final data reduction scheme some information transfer is required between these tasks and the window combining task. The topics that are closely related include double/multiple stars, photometry, imaging, spectroscopy and global astrometry.

## REFERENCES

***Babusiaux C. et al.**, 2001, The GAIA instrument and basic image simulator, GAIA technical report GAIA-CB-01*
***Bastian U.**, 2003, Reference Systems, Conventions and Notations for Gaia, GAIA technical report GAIA-ARI-BAS-003 Version 2.0*
***Dollet C. and Bijaoui A.**, 2003, Report on the combination of ASM patches, GAIA technical report IC-CDAB-01*
***Fruchter A.S. and Hook R.N.**, 2002, PASP, 114, 144*
***Gilliland R.L., Nugent P.E., Phillips M.M.**, 1999, ApJ, 521, 30*
***Høg E. et al.**, 2003, Scientific requirements for the on-board processing, GAIA technical report GAIA-CUO-117*

## Acknowledgements

*Frédéric Arenou, Anthony Brown, Erik Høg and Henri Boffin are acknowledged for all the discussions and ideas related to this work. The author wants to thank the ESA PRODEX "From Hipparcos to Gaia: Double Stars" 14847 /00/NL/SFe (IC) for the financial support of this work.*