Internship report

Made by Georgy Rozhkov, EXUL employee

The following tasks were done during the internship in South Korea, Seoul on August 18-31, 2012:

- The following lectures of IEU postdocs and leading specialists were listened:
  - George Smoot:
    1. IEU & International Research Perspective
    2. Cosmic Microwave Background (CMB)
  - Bruce Grossan:
    1. IEU and International Research
    2. Gamma-Ray Bursts Astrophysics at EWHA
    3. UBAT
    4. From Emission to Detectors to Data: Natural View, Detector View, Scientific View
    5. UFFO Slewing Mirror Telescope
    6. Intro to Optical Measurements
    7. BigBOSS
    8. Just Fibers
  - Arman Shafieloo
    1. Probing cosmological models using modern statistical technologies
  - Stephen Appleby
    1. Cosmology and Dark Energy
  - Teppei Okumura
    1. Large scale structures of the Universe
  - Frederico Arroja
    1. Introduction to Astrophysics
  - Jakub Ripa
    1. General talk about astrophysics
- The lectures of accompanying prof. Galkin were listened:
  1. Simulations of physical events in modern programming medium
  2. BDRG & GRBs
  3. Nanodosimetry approach to single event effects
  4. Spatial-angular distribution of Cherenkov's radiation as a key to primary cosmic rays with energy $E \sim 10^{15} \div 10^{17}$ eV
- **The data analysis project** (was done partly)

  Main goal of our project: to study the real data from SWIFT satellite BAT with a view to find out the minimum size of similar satellite, which is able to make some statistics on GRBs.

  My tasks:

  1. To develop computer program works like **trigger** and able to detect bursts

  Main criterion for burst detection: 5σ deviation from the average level, where σ is standard deviation of the sample and calculated by the formula (1)

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}. \quad , \qquad (1)$$

  where n is number of bins in sample, $\bar{x}$ is average signal.

What I have is 65 bursts data from http://heasarc.gsfc.nasa.gov/cgi-bin/W3Browse/swift.pl
Here is the list of my bursts: 120328A, 120327A, 120326A, 120324A, 120320A, 120312A, 120311B, 120311A, 120308A, 120305A, 120224A, 120219A, 120215A, 120213A, 120212A, 120211A, 120121A, 120119A, 120118B, 120116A, 120106A, 120102A, 111229A, 111228A, 111225A, 111215A, 111212A, 111210A, 111209A, 111204A, 111129A, 111123A, 111121A,. 110422A, 110420B, 110420A, 110414A, 110411A, 110402A, 110328A, 110319A, 110318B, 110315A, 110312A, 110305A, 110223B, 110223A, 110213A, 110212A, 110210A, 110208A, 110205A, 110201A, 110128A,110119A, 110112A, 110106B, 110106A, 110102A, 101225A, 101224A, 101219B, 101219A, 101213A, 101117B.

Main characteristics of these files are:

Table 1: Bursts input data files properties

Time resolution	0.064s
Total time duration of data stream	$10^3 \div 10^5$s
Event presence in data stream	Yes, mostly
Input data format	.lc.gz
Number of energy channels	4, used only 1,2 and 3
Prehistory of events	Usually 800÷1000 s

I downloaded these 33 .lc.gz files and then converted them to .txt files with “fv” program, which was recommended by Bruce Grossan.
Here is the example of 110102A GRB data in fv program, plot shows sum of counts on first three channels on OY and time in seconds after beginning of data stream on OX (Fig.1 and Fig.2)

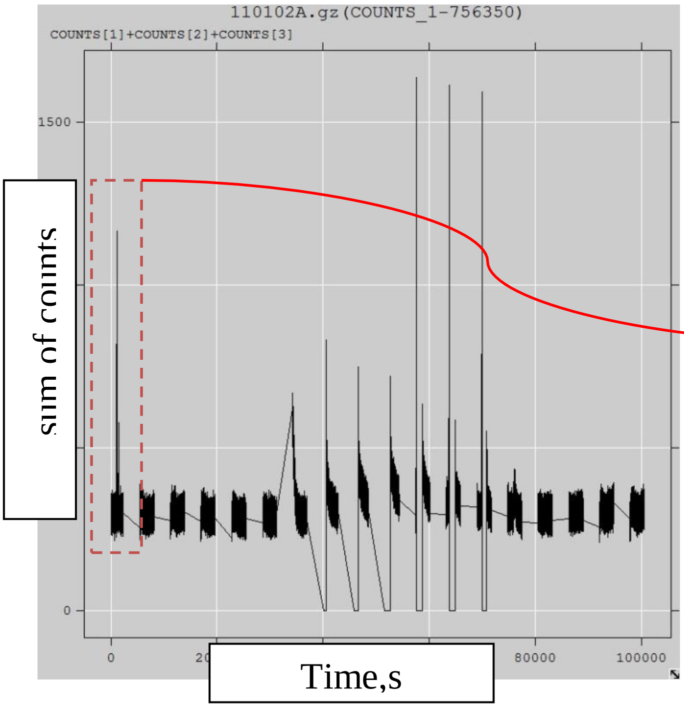

Fig.1: Full data file example (“fv” plot)

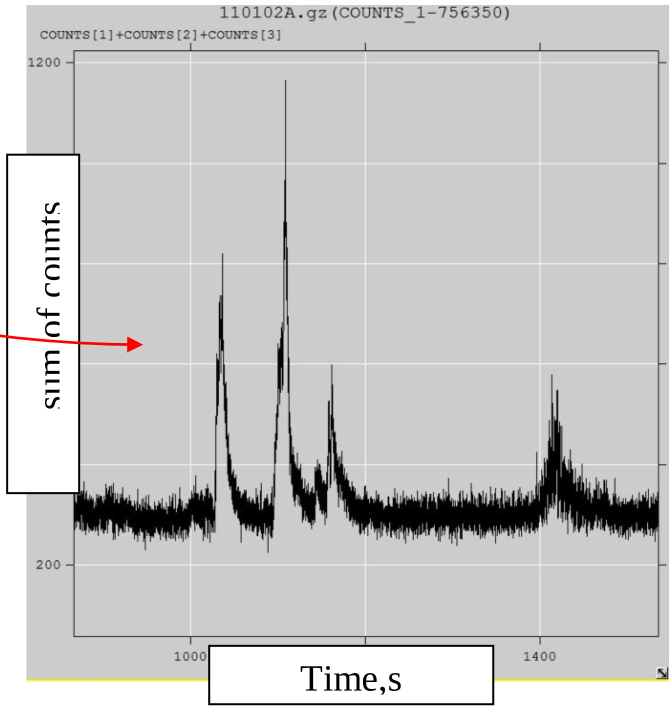

Fig.2:Event part example (“fv” plot)

The principle of my program:

- Firstly, the data array of 6250 bins (400s in time) is loaded to program from .txt file made by "fv" program
- Secondly, counts on firsts three channels are summed (4th channel is not in our interest)
- Thirdly, the average level of background is calculated for all of this 6250 bins. It makes sense, because we know exactly, that there is about 800 to 1000 s of prehistory (no events!).
- Fourthly, we act the signal by subtracting the average background level
- Then, σ is calculated by the formula (1) with n=6250

This formulae works quite good because of random nature of background fluctuations, where values are independent and normally distributed (it was checked by Kyrill Saleev).

- After that I have to use 4 different time durations for trigger windows: 0.5, 1.5, 2, 4 s for detection of GRBs with the variety of lasting: from 0.5s and higher.

Each window slides from one successive bin to the next one and estimates the average level of signal (which is determined as total number of counts minus background) in the whole time window, and if it is higher than 5σ – trigger works and writes down the time of burst, σ, Signal and noise values in output .txt file. To estimate the noise value in each window I use the following formula (2):

$$\sigma_{window} = \sigma * \sqrt{N-1}, \qquad (2)$$

where N is number of bins in each time window (which is 8, 24, 32, 64 bins respectively).

- After the program is done with this 6250 bins array, the next one array is loaded from the same data .txt file, and so on until total time of data loaded is equal to 3200 seconds (which is enough to find the burst and to equal to the average duration of one curcuit of satellite). One should mention, that σ and σ_{window} values are calculated only one time from the beginning of the procedure for each burst, and these quantities remain in internal memory of the program until the procedure is done.
- Eventually, on proceeding each of .txt files with BAT data of bursts, we get an output .txt file with the list of names of bursts, and detected ones list for each of time windows:

Table 2: Burst output .txt file example

GRB 120328A ch1+ch2+ch3			
№ of detector	Average signal, counts	Noise, counts	Time window,s
Trigger №1	498	55.72	0.512
Trigger №2	482	105.71	1.536
Trigger №3	471	123.38	2.048
Trigger №4	450	177.03	4.096

So, the first task was done and I got the .txt file with the list of my bursts and their characteristics. Unfortunately, some of files were not proceeded due to the following reasons:

- False triggering: SWIFT trigger worked, but no reasonable and observable burst can be seen by eyes. Such triggering can be caused by some different criterion used in BAT and unknown to me (exactly not "5σ") or just a technical error.
- Sharp and fast drops in background: burst or just anomaly?
- Another thing connected with long duration high amplitude changes in background: How one can be sure, that he observes GRB, but not high-background area?

Contradicting ones: 112115A, 111204A, 111212A,120221A, 120219A, 120311B and some others.

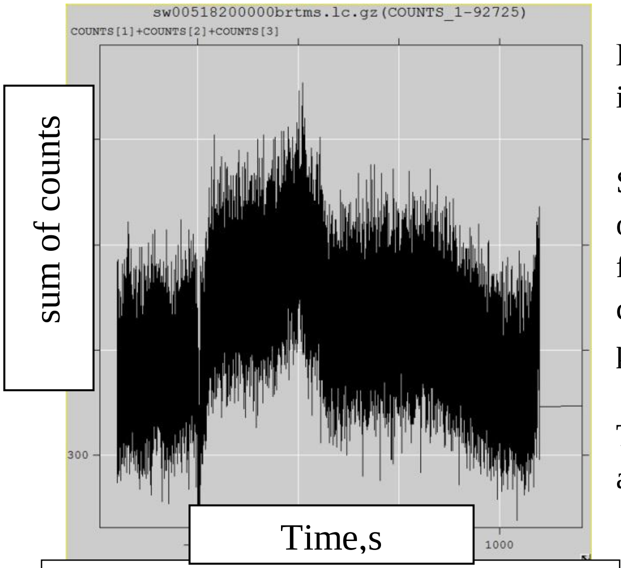

Fig.3: Example,"fv" plot of strong magnitude changes in background

Here is an example of "fv" plot of strong amplitude changes in background, but BAT claims it to be a burst (Fig.3)

So, my point is that here we can observe 3 or even 4 periods of different background levels with very high and strong fluctuations. Obviously, the situation with detection in this case is not so easy even for a BAT trigger, but for my program – it's impossible.

Thus, most of data files were proceeded (53 for 0.5s trigger and 65 for other ones), and first task was done.

Let me now turn to the next task:

2. **My second task** was to find the maximum signal to noise ratio for each of bursts studied with a view to make statistics of S/N distribution of SWIFT BAT detected GRBs and, as was said above, to find out the minimum size of similar satellite, which is able to make some statistics on GRBs (it must detect ~30 GRBs/year) by making a graph of a ratio Nx/Nswift per year and Ax/Axwift, where x is hypothetic satellite with the similar equipment, but smaller effective detection area A, N is number of GRBs, detected per year . We were expecting $(Nx/Nswift) \sim (Ax/Axwift)^{1/2}$

- But now we should consider the signal in each burst as integrated flux over the whole trigger window, and the noise value meaning remains the same. As each window slides through the whole data array, we must find the maximum integrated flux value among all others (it's obviously in burst time). So, we have formula to calculate signal to noise ratio (3):

$$(S/N)_{max} = \left(\overline{S_{window}}\right)_{max} * N/\sigma_{window}, \qquad (3)$$

where $\overline{S_{window}}$ is averaged signal (in normal sense) in each window, N remains the number of bins in window and σ_{window} is window noise. I used old values of σ_{window}, which were calculated earlier in my first task, and recalculated new signal, using its definition.

- As result of this task I got .txt file with the values of S/N ratio for each processed burst for 4 triggers time windows, and it looks like that:

Table 3: Bursts output .txt file example 2

Quantity	S/N 1	S/N 2	S/N 3	S/N 4
Trigger Time	0.512s	1.536s	2.048s	4.096s
Values	65,71451	105,1344	120,4239	163,5565
	88,95947	140,693	160,9821	216,5792
	178,5898	267,5409	294,5137	350,5845

- After that I import this files in OriginPro 8 program to make some histograms of S/N distribution for all 4 trigger windows used. And after scaling and exporting graphs to .jpeg or .bmp files, here is the result of my 2 week work:

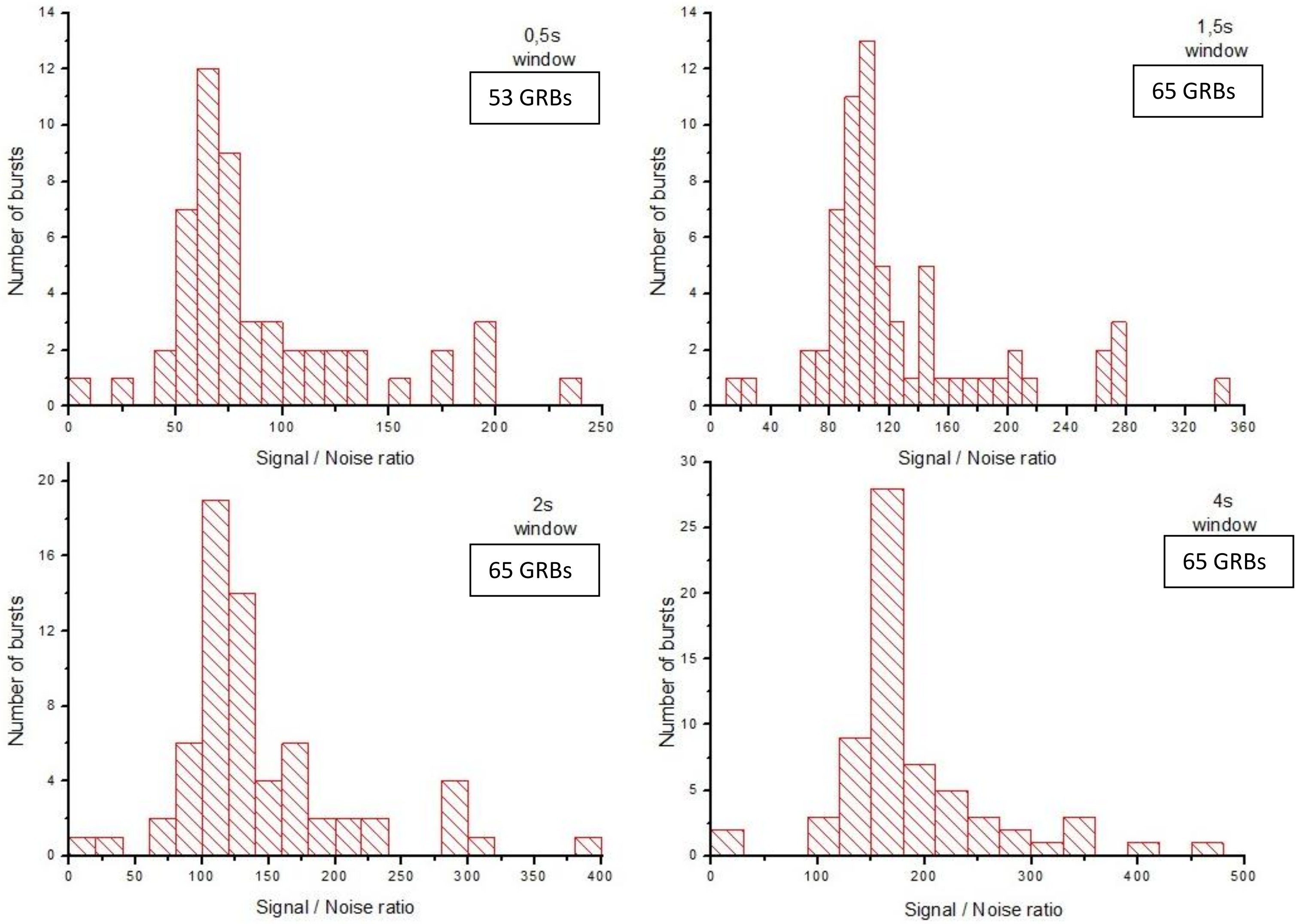

Fig.4: Histograms of Signal/Noise ratio distribution for 65(53) GRBs and 4 time windows of triggering. Graphs made by OriginPro 8.

We can see in Fig.4 the displacement of peak to the higher values of S/N ratio with the growth of time window, and this is very good up to the 2s upper limit – after that we will loose most of short GRBs and some of formally long GRBs too, trying to apply such a trigger window. I must notice that we studied mostly LGRBs with duration about 5 to 30s, so the fall in efficiency of detecting SRGBs was not considered.

- But the last part of our data analysis project was not done because of lack of time, so we didn't found relation between Nx/Nswift per year and Ax/Axwift and didn't made graph of it.

This is it about our data project and my work on it, let me now turn to other points of our internship:

- We visited and examined IEU optical laboratory, in which at the moment carrying out some experiments with optical fibers to be used in BigBOSS project:
  1. Measure-Collimated Light Test
  2. Incoherent Beam test
  3. Rotational durability test
- I took part in carrying out first and second of these experiments, third one was not done because of lack of time. **Main goal of experiments**: to study the radial distribution of output beam intensity.

<u>My task</u> was to write the program able to process input pictures, and then I had to make graphs of radial distribution of intensity. To be more precise:

1. <u>Measure-Collimated test</u>

Here is the idea of experiment, it was shown in Bruce Grossan's lecture he gave us in IEU (Fig.5)

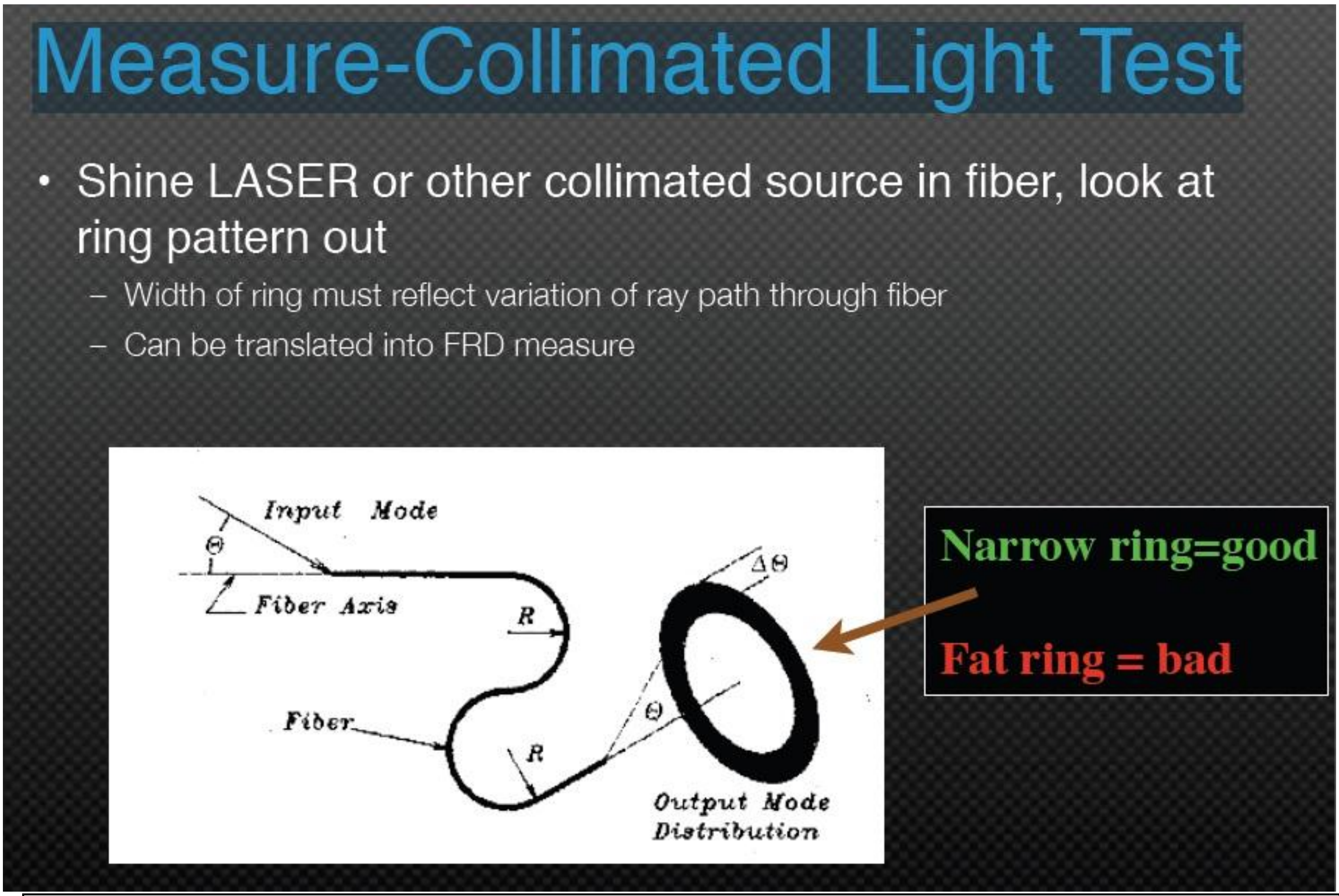

Fig.5: The idea of measure-collimated test, picture taken from Bruce Grossan's presentation in IEU

So, our **experimental setup** was just a laser, pointed at the fiber at some known angle, and CCD-matrix, which read the output mode distribution and turned it into picture with the analog-to-digital converter. Picture can be captured and saved on PC in .pgm format. We made pictures for 2 fibers with 2 angles (23˚ and 36˚) and repeated these measurements 3 times to have some statistics and avoid casual mistake.

What I have as input material is 18 .pgm pictures, looks like this (Fig.6):

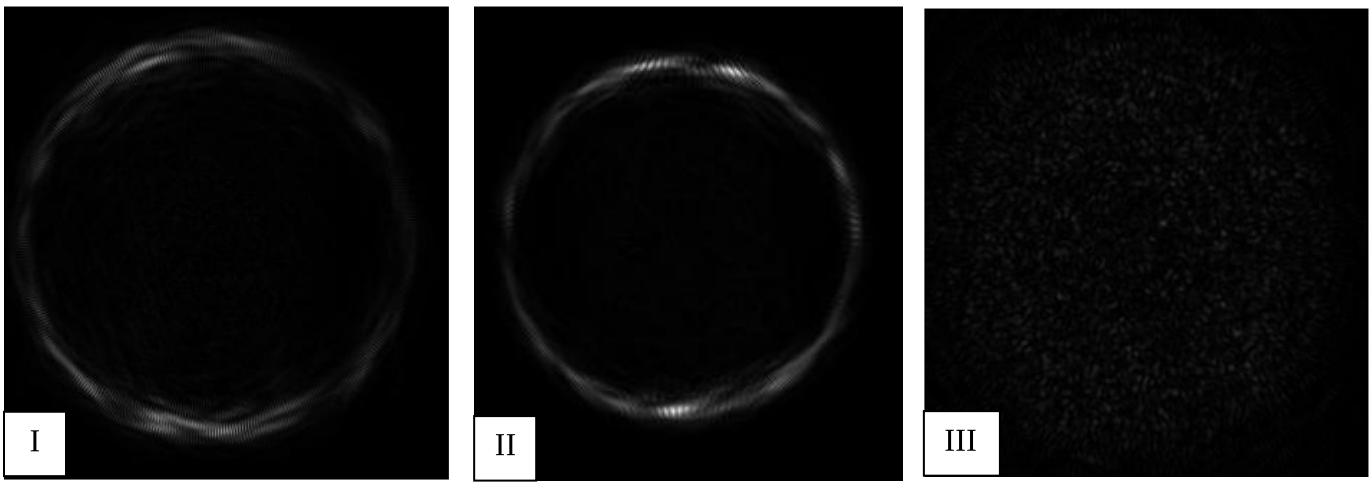

Fig.6: Example of input data files: pictures were compressed and only significant part is shown. As we can see, intensities of I and II samples are very different, and III is example of bad and blurred distribution.

2. <u>Incoherent Beam test</u>

Here is the scheme of **experimental setup** (Fig.7):

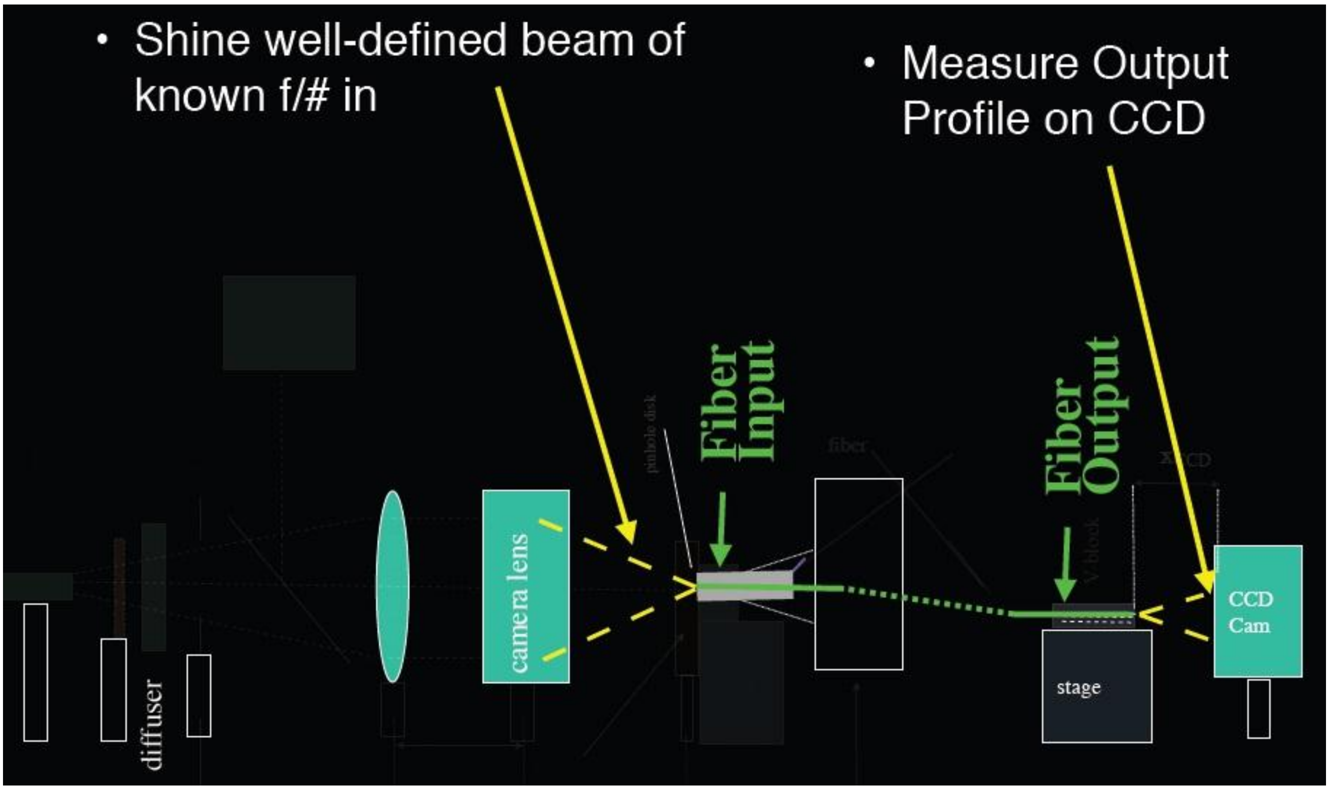

Fig.7: The idea of incoherent beam test, picture taken from Bruce Grossan's presentation in IEU

Thus, the main **idea of this experiment** – to compare input and output beam profiles in size and shape and find out if they are correspond with each other. We carried measurements with 2 fibers with 2 apertures of camera lens: f8 and f16, then repeated these 3 times.

What I have as input material is 12 .pgm pictures, looks like this (Fig.8):

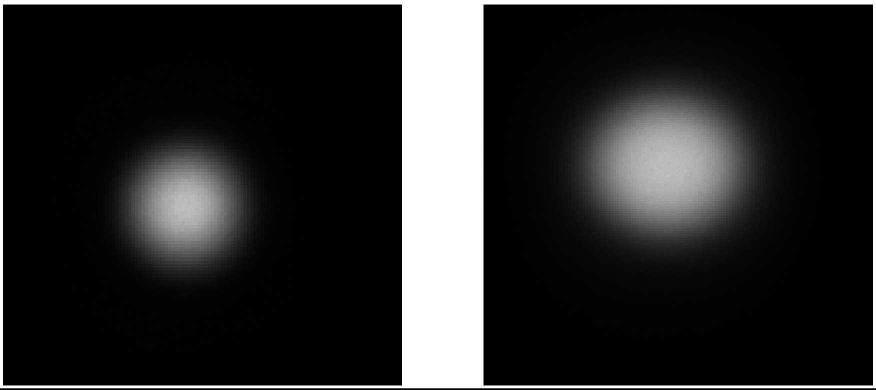

Fig.8: Example of input data files: pictures were compressed and only significant part is shown

The main characteristics of input data files:

Table 4: Fibers input files properties

Pictures format	.pgm
Number of samples	30
Picture resolution	1280x960 pix^2
Color depth	8 bits per pixel

Firstly I convert input files to .bmp format using the "ImageJ" program, which was recommended by Bruce Grossan. Then I wrote computer program to process this pictures and get the radial distribution of intensity data.

Main principle of my program for processing input pictures:

- Input .bmp file is loaded into the program as BITMAP array with the size of picture (1280x960)
- I don't know the exact position of the output beam center, so I have to find it. For this purpose I used the common formula of center of weight (intensity is implied as weight) which is (4):

$$\vec{r}_c = \frac{\sum_i \vec{r}_i m_i}{\sum_i m_i}, \qquad (4)$$

where $\vec{r_c} = (x_c, y_c)$ is radius vector of center of weight, $\vec{r_i} = (x_i, y_i)$ is radius vector of each point, m_i is an intensity of each point. I found x_c and y_c separately.

- After the center found, I have to calculate radius of each point from the central point by the (5):

$$r_i' = |\vec{r_i} - \vec{r_c}| = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}, \qquad (5)$$

where r_i' is distance of each point from a central point.

- At this stage I need just to sort all the array of points in ascending radius order. I used kind of bubble sort for 2D array of data, which is very simple, but not very effective – for an array of 150000 points it takes about 3 minutes to sort, but this method can easily be improved.
- Eventually, program writes down .txt file with the tab:

Table 5: Pictures output .txt example

Radius,pix	Intensity,units
0	191
1	186
1.41	187
2	186
2.24	184

- Then I imported these .txt files into OriginPro 8 and made graphs of radial profile of each measurement, and here are the examples of input picture and output profile (Fig.9, Fig.10):

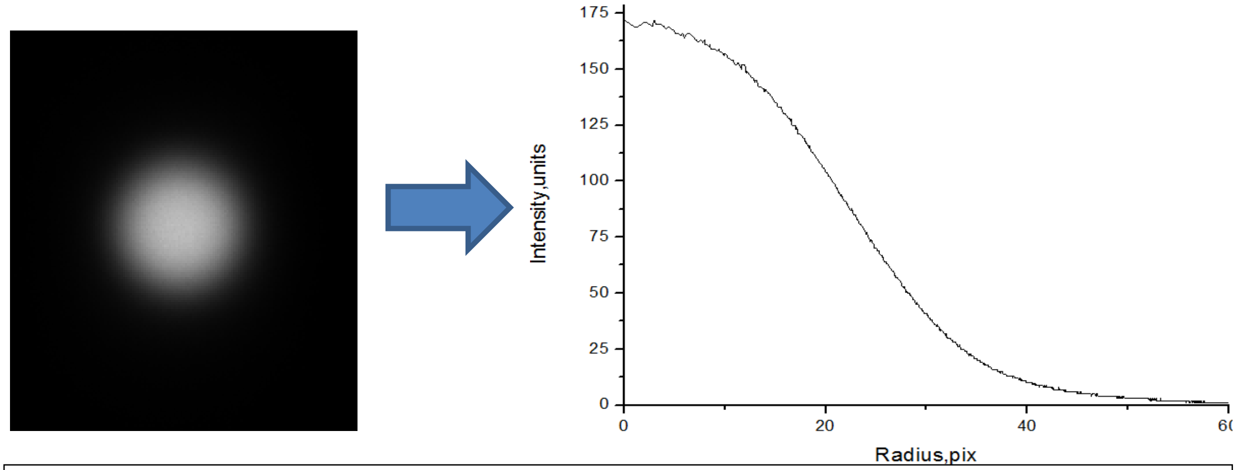

Fig.9: Example of an input picture and output radial profile plot for incoherence beam test

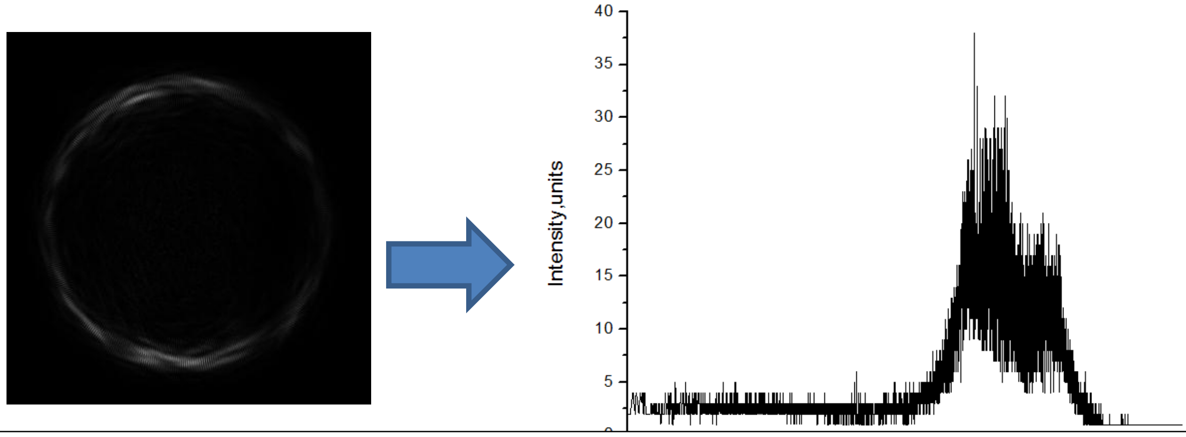

Fig.10: Example of an input picture and output radial profile plot for measure-collimated light test

Not all of our measurements were successful (or we just took bad fibers, who knows), and here is profile for highly blurred and faded picture (Fig.11)

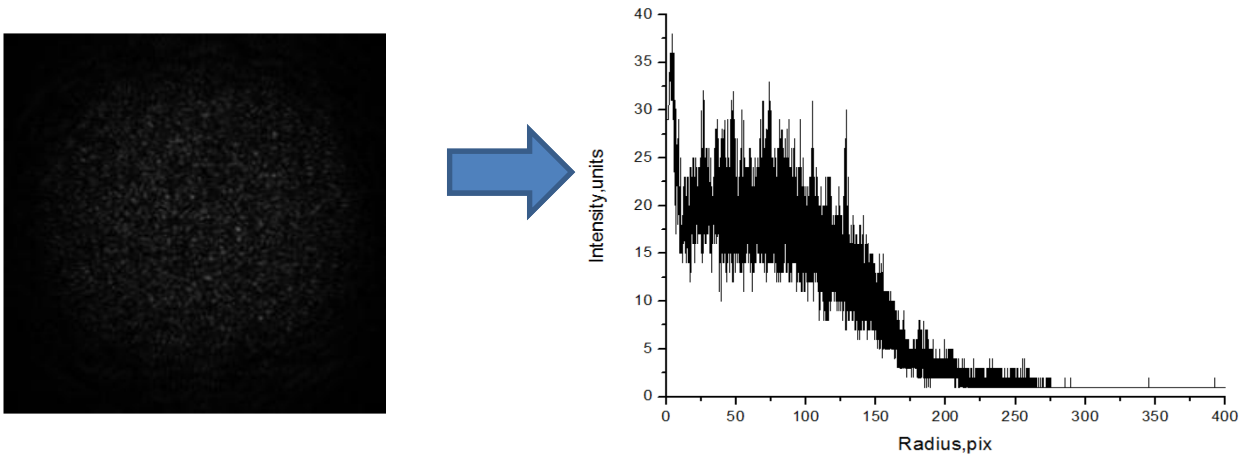

Fig.11: Example of bad input picture and output radial profile plot for measure-collimated light test (it shall be the annulus, but what we see is just a cloud of dots – bad fiber probably)

So, I have finished this task in time and processed 29 pictures (one of 30 was spoiled by us), and **what I got is 29 profile plots** of radial distribution of intensity of output beam for 2 experiments made.

This is all about the experimental part; let me now turn to the last part of our internship:

- We inspected clean and "ultra-clean" rooms of IEU laboratory; there is hermetic gateway presented and powerful dust blowing in it. There are also two stabilized optical tables, temperature and humidity control systems, and other equipment for testing and assembling experimental devices.
- But, unfortunately, our tasks on acquaintance with the newest electronics (which is multichannel plate of pre-multiplier, SPAD- and LSO-detectors) and work of laboratory of assembling and testing the devices were not done. Professor Il Pak and his team did not take part in our internship, so we just had no possibility to perform this tasks.

Let me make a **conclusion** about my experience in this internship.

As a result of this internship I got a lot of new knowledge in astrophysics, insight on real research work and analysis of experimental data, work in “clean room”, got skills in programming and performing of experiments with optical fibers using the CCD-matrix camera, collimation and calibration of optical devices.

The lectures have been listened elucidated existing situation in such areas of astrophysics, as γ-ray bursts detection, optical observations, theoretical modeling, simulations of physical events and phenomena, assembly and testing of real devices, analysis of experimental data, Early Universe history and so on…

Thus I was happy to take part in this wonderful internship, and such an experience is undoubtedly very useful for me and my future scientific life.