

Astronomical Metadata Mining from FITS Files by the Telescope Software

Sergii Khlamov^a, Tetiana Trunova^a, Iryna Tabakova^a

^a Kharkiv National University of Radio Electronics, Nauki avenue 14, Kharkiv, 61166, Ukraine

Abstract

In this paper we presented the realization of the data mining approach related to the metadata of astronomical files from the big archives. Each astronomical file has the commonly defined structure, which contains the especial format of the metadata. Such metadata contain the necessary astronomical information, which is required for the proper storing, data mining, processing, analyzing under research. This realization was implemented as software called "Telescope" using the C# programming language, .NET platform, Windows Forms technology and equipped with the MDB database file for Microsoft Access DBMS. The software has two modes: console mode for the automated integration with the processing pipelines and mode with a graphical user interface (GUI) for the visualization of processing and the additional useful features. The Telescope software was designed for mining the big astronomical data from the different archives, parsing the metadata from each astronomical file, and collecting it with the further insertion into the database. Such parsed data were used for the different purposes of the astronomical image processing and machine vision. The Telescope software was developed during research under the CoLiTec project and was tested with the astronomical files from several archives on the different observatories. Also, the Telescope software was successfully implemented and installed on the astronomical image processing pipelines in such observatories.

Keywords

Data mining, big data, metadata, database, image processing, machine vision, C#, .NET, Windows Forms, Microsoft Access

1. Introduction

Almost all astronomical frames are made by the CCD-camera [1] and can be received from the different sources: archives, servers, predefined series of frames, Virtual Observatories [2], clusters, etc. Each software for preparing the astronomical frames creates them as the digital files in the FITS (Flexible Image Transport System) format [3].

This format is a digital files format for storing and transferring of their image and metadata (spreadsheets). Metadata is a kind of data, which provides the information about other data, except the original data content. There are a lot of different types of metadata, as following [4]:

- *Descriptive* metadata is an information about the resource, which is used for identification and includes the elements, like author, title, abstract, and keywords.
- *Structural* metadata is an information about the data containers and how the objects are collected. It includes the elements, like relationships, versions, types, etc.
- *Administrative* metadata is an information for managing resources (creation date, edition date, permissions, etc.).
- *Reference* metadata is an information about the static data, references to them, and contents.
- *Statistical* metadata is an information about the processes for collection, producing, and publishing the statistical data.

ICST-2023: Information Control Systems & Technologies, September 21-23, 2023, Odesa, Ukraine

EMAIL: sergii.khlamov@gmail.com (S. Khlamov); tetiana.trunova@nure.ua (T. Trunova); iryna.tabakova@nure.ua (I. Tabakova)

ORCID: 0000-0001-9434-1081 (S. Khlamov); 0000-0003-2689-2679 (T. Trunova); 0000-0001-6629-4927 (I. Tabakova)



© 2023 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

- *Legal* metadata is a legal information about the copyright, creator, and licensing.

The main purpose of metadata is to provide an information about the different aspects of original data and to summarize a basic information about it, which make tracking and processing it easier. The examples of metadata are as following: time and date of data creation, its meaning and purpose, creator or author, location, file size, used standards, sources, quality, etc.

For example, the digital image includes the metadata, which describes image size, its color depth, resolution, creation time and date, exposure time, etc. A metadata of the text document includes an information about author, processing time of document, short summary, etc. The web pages includes metadata, which describes a description of page content, and keywords linked to it.

In astronomy metadata is used for the different image processing and machine vision purposes [5], like analyzing, acquiring, pre-processing, processing, and extraction of high-dimensional astronomical information [6].

Such purposes [7] are focused on but not limited to the following tasks: brightness equalization and background alignment [8], object's images detection [9], moving objects detection [10], astrometry of object's image [11], photometry of object's image [12], the estimation of the object's image and motion parameters [13], reference objects cataloging [14], objects recognition [15], time series analysis [16], Wavelet coherence analysis [17] and others.

There are different types of astrophysical objects that can be detected, like galaxies, stars, robots [18, 19], drones [20], rockets, satellites [21], and even comets or asteroids [22].

In this paper we presented a description of the astronomical metadata from the real examples of CCD-images [23], usage in the data mining approach from the big archives, and its implementation as a developed *Telescope* software, which is designed for mining the big astronomical data from the different archives, parsing the metadata from each astronomical file, and collecting it with the further insertion into the database.

2. Big astronomical metadata

Almost all astronomical frames have a FITS format with standardized file structure and extension. In the common case the astronomical file extensions are: *.fits, *.FITS, *.fts, *.FTS, *.fit, *.FIT. Such FITS format is commonly used for the transformation, transferring, and archiving of astronomical data.

The FITS format was developed by National Aeronautics and Space Administration (NASA) and is accepted as an international astronomical standard and is used by many astronomical and scientific organizations, like International Astronomical Union (IAU) [24], and other national and international organizations that deal with the astronomy or related scientific fields.

The FITS format is commonly used for the storing the data without the image, like spectrums, photons list, data cubes or even structured data, such as databases with multiple tables. The FITS format includes many provisions to describe the photometric and spatial calibration, as well as image metadata.

The structure of the FITS file consists of a header with metadata and a binary image. The header size is 2880 bytes and contain the list of human readable metadata in fixed string form of 80 symbols. Each string is an ASCII [25] stroke, which contains the pair with key and value, and have the common form: "*KEYNAME = value / comment string*". Each header block should be ended with the especial key "*END*" with the empty value. The example of a header with metadata of the real astronomical FITS file is presented in the Figure 1.

There are the minimum list of the required keywords to make the header and the whole FITS file valid. They are:

- "*SIMPLE*" (file conforms to FITS standard);
- "*BITPIX*" (bitrade of FITS file, bits per pixel);
- "*NAXIS*" (number of axes);
- "*NAXIS1*" (number of points along axe 1);
- "*NAXIS2*" (number of points along axe 2);
- "*END*".

```

SIMPLE = T
BITPIX = 16 /8 unsigned int, 16 & 32 int, -32 & -64 real
NAXIS = 2 /number of axes
NAXIS1 = 512 /fastest changing axis
NAXIS2 = 512 /next to fastest changing axis
BSCALE = 1.0000000000000000 /physical = BZERO + BSCALE*array_value
BZERO = 32768.000000000000 /physical = BZERO + BSCALE*array_value
DATE-OBS= '2017-03-30T21:39:16' /YYYY-MM-DDThh:mm:ss observation start, UT
EXPTIME = 60.000000000000000 /Exposure time in seconds
EXPOSURE= 60.000000000000000 /Exposure time in seconds
SET-TEMP= -40.000000000000000 /CCD temperature setpoint in C
CCD-TEMP= -40.062500000000000 /CCD temperature at start of exposure in C
XPISZ = 48.000000000000000 /Pixel Width in microns (after binning)
YPISZ = 48.000000000000000 /Pixel Height in microns (after binning)
XBINNING= 2 /Binning factor in width
YBINNING= 2 /Binning factor in height
XORGSUBF= 0 /Subframe X position in binned pixels
YORGSUBF= 0 /Subframe Y position in binned pixels
READOUTM= '1 MPPS ' / Readout mode of image
FILTER = 'V ' / Filter used when taking image
IMAGETYP= 'Light Frame' / Type of image
OBJECT = 'DO Dra '
OBJCTRA = '11 43 38' / Nominal Right Ascension of center of image
OBJCTDEC= '+71 41 20' / Nominal Declination of center of image
OBJCTALT= ' 67.3413' / Nominal altitude of center of image
OBJCTAZ = ' 0.2784' / Nominal azimuth of center of image
OBJCTHA = ' -0.0226' / Nominal hour angle of center of image
SITELAT = '48 56 06' / Latitude of the imaging location
SITELONG= '22 16 27' / Longitude of the imaging location
JD = 2457843.4022685187 /Julian Date at start of exposure
JD-HELIO= 2457843.4039910869 /Heliocentric Julian Date at exposure midpoint
AIRMASS = 1.0833536254388081 /Relative optical path length through atmosphere
FOCALLEN= 9000.0000000000000 /Focal length of telescope in mm
APT DIA = 1000.0000000000000 /Aperture diameter of telescope in mm
APTAREA = 777544.20340061188 /Aperture area of telescope in mm^2
SWCREATE= 'MaxIm DL Version 5.12' /Name of software that created the image
SBSTDVER= 'SBFITSEXT Version 1.0' /Version of SBFITSEXT standard in effect
TELESCOP= 'VNT ' / telescope used to acquire this image
INSTRUME= 'FLI ' / instrument or camera used
OBSERVER= 'DPV '
NOTES = ' '
FLIPSTAT= ' '
SWOWNER = 'Amigo ' / Licensed owner of software
END

```

Figure 1: The example of a header with metadata of the real astronomical FITS file

3. Telescope software

Under the research in scope of the *CoLiTec* project [26] we have developed the *Telescope* software for mining the big astronomical data from the different storages and archives, parsing the metadata from each astronomical file, and collecting it with the further insertion into the database. Such parsed data were used for the different purposes of the astronomical image processing and machine vision.

3.1. Technical implementation

The *Telescope* software realized the different data mining tasks, like receiving, storing, selecting, preprocessing, transforming, useful data extraction, classification, and knowledge discovery in databases (KDD) [27]. The following stack of technologies were used for the software development: C# programming language, .NET platform, Windows Forms technology and MDB database file for Microsoft Access database management system (DBMS). The Windows Forms (WinForms) is an open-

source and free graphical library, which is in a scope of the Microsoft .NET Framework [28] and play the role as a platform for developing the client applications for desktop, laptop, and tablet PCs.

As a database the developers have selected a MDB database file as a native format for the Microsoft Access DBMS with “.mdb” extension [29]. There are different fields located in several tables of the MDB database file in the *Telescope* software that represent the necessary for research metadata of the astronomical file. The list of such fields with their description and types including the system fields are presented in the Table 1.

Table 1

Metadata in header of the FITS file

Field	Type	Description
FitCode	Long Text	Unique code of each FITS file
FitName	Short Text	Name of FITS file on storage
FitPath	Short Text	Full path to FITS file on storage
RA	Number	Right ascension of frame center
DE	Number	Declination of frame center
Date	Date	Date of observation
Time	Time	Time of observation
Exp	Number	Exposure time
BitPix	Number	Pixel bit rate
Height	Number	Height of frame
Width	Number	Width of frame
PixHeight	Number	Height of pixel
PixWidth	Number	Width of pixel
Focus	Number	Focal length of telescope
Apertrure	Number	Aperture size of telescope
Temp	Number	Temperature of CCD-camera
Long	Number	Longitude of telescope
Lat	Number	Latitude of telescope
Alt	Number	Altitude of telescope
Instrum	Short Text	Instrument name
Telescope	Short Text	Telescope name
Observer	Short Text	Observer name
Obj	Short Text	Investigated object name

The *Telescope* software has two different modes: console mode for the silent automated integration with the image processing pipelines and mode with a graphical user interface (GUI) for the visualization of processing and the additional useful features.

3.2. Console mode

The console mode of the *Telescope* software is designed for the integration with processing pipeline as a precondition step for searching, collecting, and parsing the astronomical metadata from the files in archives. The *Telescope* software in console mode can be launched in the Windows Command Prompt using the following list of the implemented commands for:

- showing help information with all available commands (“*help*”);
- adding the new astronomical FITS file with metadata to be parsed and inserted into the database (“*add [PathToFile.fits]*”);
- adding the new already parsed astronomical metadata from the text file into the database (“*add [PathToFile.txt]*”);
- searching for the metadata by the different optional criteria in the database (“*find [Observer] [Telescope] [RA] [DE] [Date] [Time] [PathToFile.txt]*”);

- searching for the metadata in a range by the especial criteria in the database (“*find [RAfrom] [RAto] [DEfrom] [DEto] [PathToFile.txt]*”);
- exporting the MDB database file to the user’s local folder (“*export [PathToFolder]*”);
- importing the MDB database file from the user’s local folder (“*import [PathToDBFile.mdb]*”).

3.3. Mode with GUI

The mode with GUI of the *Telescope* software is designed for the independent big astronomical metadata preparation. The processing pipeline includes the following steps:

- selecting the work folder including subfolders with the different astronomical FITS files (see Figure 2);

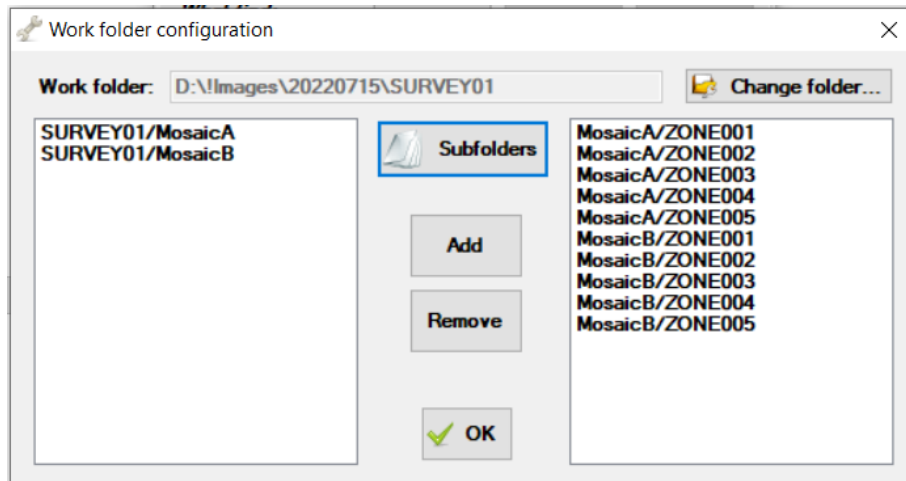


Figure 2: “Work folder configuration” window in the *Telescope* software

- recurrency searching for the astronomical FITS files according to the extensions in the work folder and forming the results list (see Figure 3 (left));

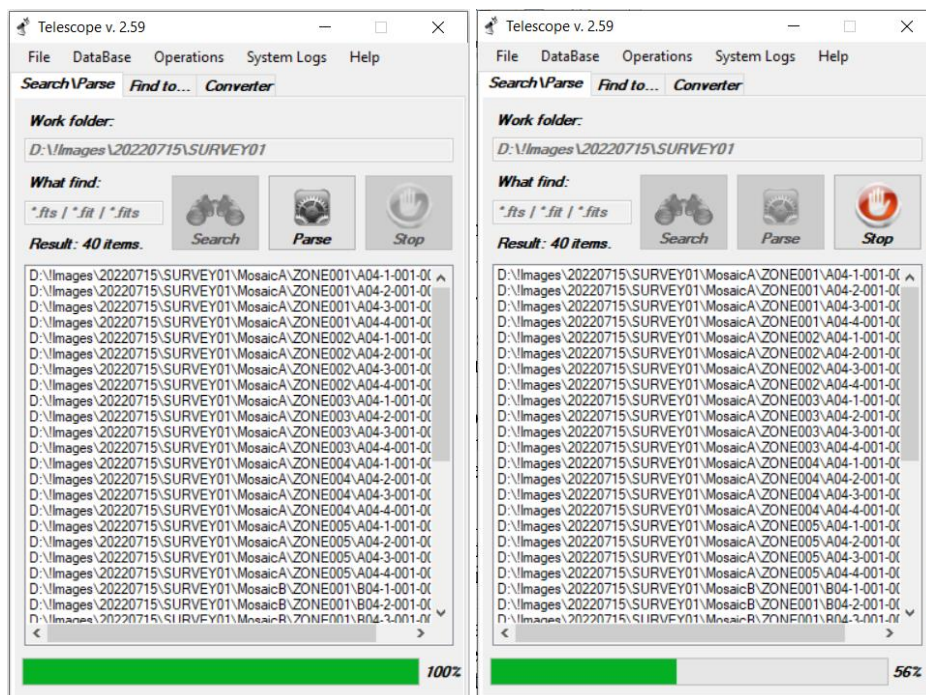


Figure 3: The “Search/Parse” page with the list of results in the *Telescope* software

- parsing the astronomical FITS files from the results list, extracting the metadata and insertion into the MDB database file (see Figure 3 (right));
- finding the different astronomical metadata in the MDB database file and forming the results list (see Figure 4 (left));

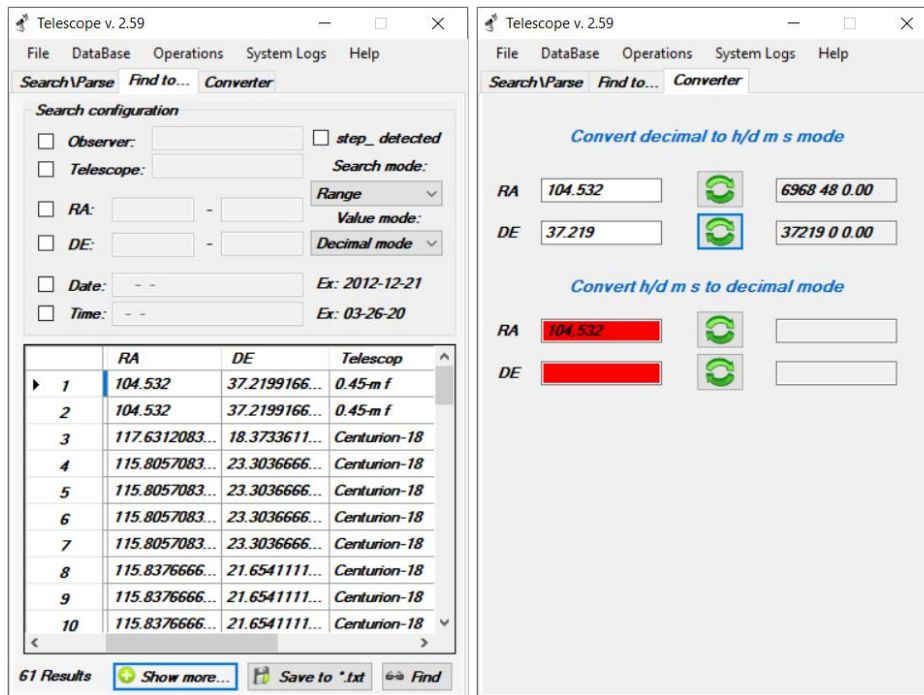


Figure 4: The “Find to” (left) and “Converter” (right) pages in the *Telescope* software

- converting from arcseconds (angular hours / minutes / seconds) to decimal (1), (2) and vice (3), (4) (see Figure 4 (right)) [11];
- displaying all available astronomical metadata in the MDB database file according to the results list (see Figure 5);

61 Units	FitName	RA	DE	Exp	BitPix	Height	Width	PixHeight	PixWidth	Focus	Aperture	Temp	Long	Lat	Alt	Instrum	Telescop	Obj
1	FCLT_PSN J065...	104.532	37.2199166666...	300	32	1024	1024	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	0.45-m f	PSN J06580768...
2	YCLT_PSN J065...	104.532	37.2199166666...	300	32	1024	1024	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	0.45-m f	PSN J06580768...
3	A05-1-001-001.fits	117.631208333...	18.3733611111...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A05-1
4	A04-1-001-001.fits	115.805708333...	23.3036666666...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A04-1
5	A04-2-001-001.fits	115.805708333...	23.3036666666...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A04-2
6	A04-3-001-001.fits	115.805708333...	23.3036666666...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A04-3
7	A04-4-001-001.fits	115.805708333...	23.3036666666...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A04-4
8	A03-1-001-001.fits	115.837666666...	21.6541111111...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A03-1
9	A03-2-001-001.fits	115.837666666...	21.6541111111...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A03-2
10	A03-3-001-001.fits	115.837666666...	21.6541111111...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A03-3
11	A03-4-001-001.fits	115.837666666...	21.6541111111...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A03-4
12	A01-1-001-001.fits	115.892833333...	18.3555555555...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A01-1
13	A02-1-001-0 01.fits	115.866583333...	20.0051944444...	240	16	3056	3056	0.012	0.012	1270	455	-24.9375	-105.52843	32.90323	2217	FLI - New	Centurion-18	A02-1
14	A02-2-001-001.fits	115.866583333...	20.0051944444...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A02-2
15	A02-3-001-001.fits	115.866583333...	20.0051944444...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A02-3
16	A02-4-001-001.fits	115.866583333...	20.0051944444...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A02-4
17	A01-2-001-001.fits	115.892833333...	18.3555555555...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A01-2
18	A01-3-001-001.fits	115.892833333...	18.3555555555...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A01-3
19	A01-4-001-001.fits	115.892833333...	18.3555555555...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A01-4
20	A02-3-001-001.fits	115.866583333...	20.0051944444...	240	16	3056	3056	0.012	0.012	1270	455	-25	-105.52843	32.90323	2217	FLI - New	Centurion-18	A02-3

Figure 5: Astronomical metadata in the database according to the results list in the *Telescope* software

- exporting/importing the MDB database file;
- logging the searching, parsing process, error handling.

In general, positional coordinates of objects can be of two types: decimal Cartesian (x and y) in the image plane and stellar (right ascension RA and declination DE) as angular coordinates in the sky [14]. So, to use the appropriate type of positional coordinates, the following mutual recalculation can be used:

$$x_d = \left(\left(\left(\frac{RA_S}{60} + RA_M \right) / 60 \right) + RA_H \right) * 15 ; \quad (1)$$

$$y_d = \left(\left(\frac{DE_S}{60} + DE_M \right) / 60 \right) + DE_H ; \quad (2)$$

$$RA_H : RA_M : RA_S = \left\{ \begin{array}{l} (int) \frac{x_d}{15} \\ (int)(x_d - RA_H) * 60 \\ (x_d - RA_H - RA_M) * 60 \end{array} \right\} ; \quad (3)$$

$$DE_H : DE_M : DE_S = \left\{ \begin{array}{l} (int)y_d \\ (int)(y_d - DE_H) * 60 \\ (y_d - DE_H - DE_M) * 60 \end{array} \right\} , \quad (4)$$

where d is a decimal value;

H, M, S are the angular hours, minutes, seconds accordingly.

3.4. Microsoft Access solution

The relationships between tables and fields in the Microsoft Access solution of the MDB database file integrated into the *Telescope* software is presented in the Figure 6.

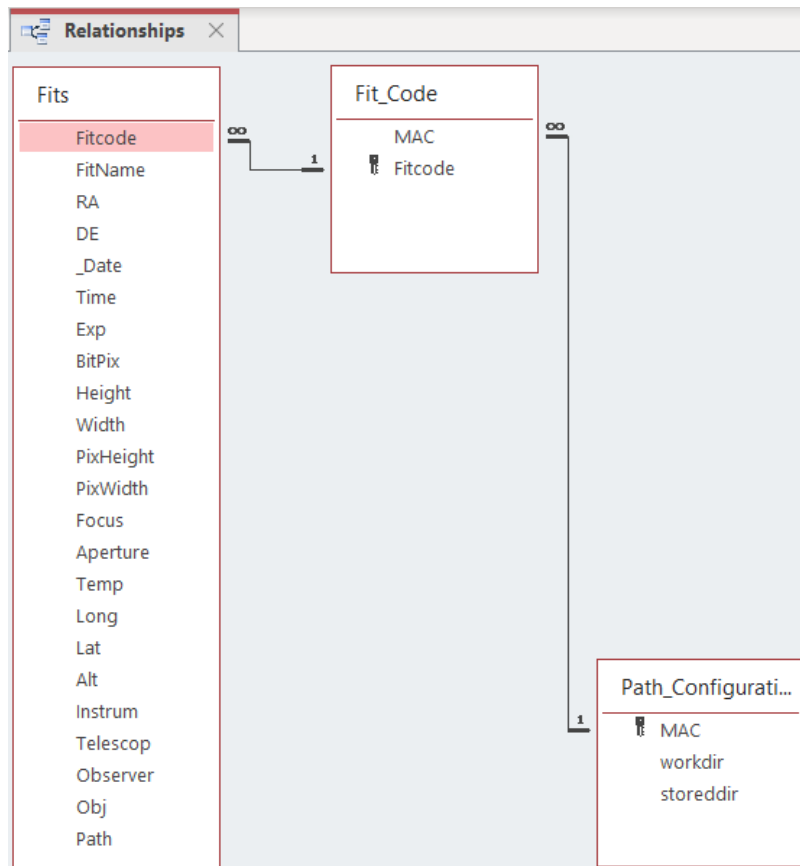


Figure 6: Relationships between tables and fields in the in the Microsoft Access solution integrated into the *Telescope* software

The example of the filled “*Fits*” table in the MDB database file integrated into the *Telescope* software with the real astronomical data is presented in the Figure 7.

Fitcode	FitName	RA	DE	Time	Exp	BitPix	Height	Width	PixHeight	PixWidth	Focus	Long	Lat
B04-3-001-001_Centurion-18	B04-3-001-001.fits	179.175291666667	-0.824583333333333	10-38-21	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
B04-4-001-001_Centurion-18	B04-4-001-001.fits	179.175291666667	-0.824583333333333	10-58-42	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A04-3-001-001_Centurion-18	A04-3-001-001.fits	115.892833333333	18.3555555555556	02-47-17	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A04-4-001-001_Centurion-18	A04-4-001-001.fits	115.892833333333	18.3555555555556	03-06-46	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A01-2-001-001_Centurion-18	A01-2-001-001.fits	115.892833333333	18.3555555555556	03-06-46	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A01-3-001-001_Centurion-18	A01-3-001-001.fits	115.892833333333	18.3555555555556	03-26-20	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A03-1-001-001_Centurion-18	A03-1-001-001.fits	115.892833333333	18.3555555555556	03-26-20	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A03-2-001-001_Centurion-18	A03-2-001-001.fits	115.892833333333	18.3555555555556	03-45-51	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A01-4-001-001_Centurion-18	A01-4-001-001.fits	115.892833333333	18.3555555555556	03-45-51	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A01-1-001-001_Centurion-18	A01-1-001-001.fits	115.892833333333	18.3555555555556	02-47-17	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A02-3-001-001s_Centurion-18	A02-3-001-001.fits	115.892833333333	18.3555555555556	02-47-17	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
FCLT_PSN J06580768+3713117	FCLT_PSN J065807	115.892833333333	18.3555555555556	02-47-17	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
YCLT_PSN J06580768+3713117	YCLT_PSN J065807	115.866583333333	20.0051944444444	02-52-07	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
FCLT_PSN J06580768+3713117	FCLT_PSN J065807	115.866583333333	20.0051944444444	02-52-07	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A02-1-001-001_Centurion-18	A02-1-001-001.fits	115.866583333333	20.0051944444444	02-52-07	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A05-1-001-001_Centurion-18	A05-1-001-001.fits	115.866583333333	20.0051944444444	03-11-37	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
YCLT_PSN J06580768+3713117	YCLT_PSN J065807	115.866583333333	20.0051944444444	03-11-37	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A02-2-001-001_Centurion-18	A02-2-001-001.fits	115.866583333333	20.0051944444444	03-11-37	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A04-4-001-001_Centurion-18	A04-4-001-001.fits	115.866583333333	20.0051944444444	03-31-11	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A03-4-001-001_Centurion-18	A03-4-001-001.fits	115.866583333333	20.0051944444444	03-31-11	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A03-1-001-001_Centurion-18	A03-1-001-001.fits	115.866583333333	20.0051944444444	03-31-11	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323
A02-3-001-001s_Centurion-18	A02-3-001-001.fits	115.866583333333	20.0051944444444	03-31-11	240	16	3056	3056	0.012	0.012	1270	-105.52843	32.90323

Figure 7: Astronomical metadata in the database according to the results list in the *Telescope* software

Each value from metadata in header of the FITS file found by the *Telescope* software was successfully parsed and filled into the appropriate field in the “*Fits*” table in the MDB database file for the further using and processing.

All interactions between such a database file and the *Telescope* software are realized by using the structured query language (SQL) as a programming language for storing and processing information in a relational database.

3.5. Metadata mining algorithm

The *Telescope* software implements the following algorithm for astronomical metadata mining.

1. Selecting the work folder including subfolders with the different astronomical FITS or TXT files.
2. Database import from the pre-filled MDB database file, which was previously exported from the last session.
3. Recurrency searching for the astronomical FITS or TXT files according to the extensions in the work folder.
4. Getting access to the astronomical FITS or TXT files (in case when location of big astronomical data is in the different remote/web archives) and download them.
5. Reading the astronomical FITS or TXT files (for FITS files splitting for two parts: header with astronomical information and body with image bytes).
6. Parsing the astronomical FITS or TXT files.
7. Astronomical data receiving from the astronomical FITS or TXT files (for FITS files from header).
8. Astronomical data converting using the different mathematical methods [30] (e.g., positional coordinates conversion).
9. Astronomical data structurization according to the patterns/classes/clusters based on the statistical modeling [31].
10. Metadata mining from the astronomical data structure according to the appropriate parameters/fields/properties.
11. Filling in the appropriate fields and tables in database using the metadata accordingly.
12. Database export to the MDB database file (if needed).

The metadata mining algorithm implemented in the *Telescope* software is presented as UML-diagram in the Figure 8.

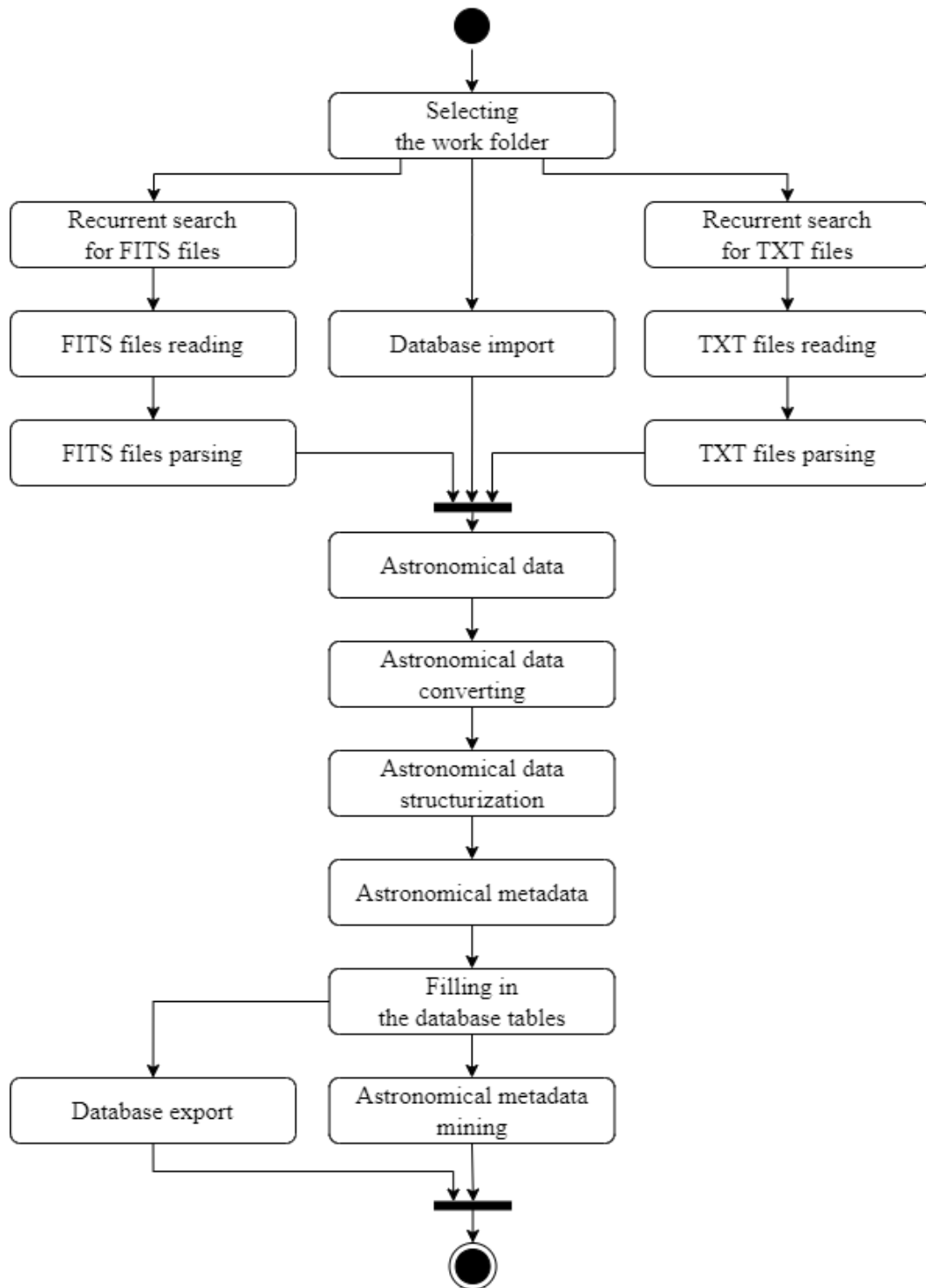


Figure 8: Metadata mining algorithm implemented in the *Telescope* software

3.6. Real astronomical examples

The *Telescope* software was installed in the different observatories (ISON-NM and ISON-Kislovodsk observatories, Vihorlat Observatory [8], Mayaki Astronomical Observatory [32, 33]), astronomical archives [2], and Ukrainian Virtual Observatory (UkrVO) [34].

The observatory "ISON-NM observatory" has the 0.4 m SANTEL-400AN telescope with CCD-camera FLI ML09000-65 (3056×3056 pixels, 12 microns).

The observatory "ISON-Kislovodsk" has the 19.2 cm wide-field GENON (VT-78) telescope with CCD-camera FLI ML09000-65 (4008×2672 pixels, 9 microns).

The observatory "Vihorlat Observatory in Humenne" has the Vihorlat National Telescope (VNT) – Kassegren telescope with 1 m main mirror with focal length 8925 mm and CCD-camera FLI PL1001E (512×512 pixels).

The Vihorlat Observatory also has the Celestron C11 telescope – Schmidt-Cassegrain telescope with 28 cm main mirror with focal length of 3060 mm and CCD-camera G2-1600 (resolution 768×512 pixels).

The Mayaki observing station of "Astronomical Observatory" Research Institute of I. I. Mechnikov Odessa National University has the 0.48 m AZT-3 telescope – reflector with focal length 2025 mm and CCD-camera Sony ICX429ALL (resolution 795×596).

Totally was processed up to 1 million astronomical files both archived and original formed from the telescopes with a lot of metadata in their headers. The received information was inserted to the database as a big data and processed by the UkrVO, which also processed a lot of different big astronomical archives [35] both digital and even plates.

4. Conclusions

The *Telescope* software with the realization of the data mining approach related to the metadata of astronomical files from the big archives was developed. The *Telescope* software is implemented using the C# programming language, .NET platform, Windows Forms technology and equipped with the MDB database file for the Microsoft Access DBMS. The SQL was used as a programming language for storing and processing information in a relational database.

The software has two modes: console mode for the automated integration with the processing pipelines and mode with a graphical user interface (GUI) for the visualization of processing and the additional useful features. The *Telescope* software was designed for mining the big astronomical data from the different archives, parsing the metadata from each astronomical file, and collecting it with the further insertion into the database. Such parsed data were used for the different purposes of the astronomical image processing and even for the Wavelet coherence analysis purposes [36].

The *Telescope* software was developed during research under the *CoLiTec* project [37]. It was tested with up to 1 million astronomical files from several archives on the different observatories. Such archives included astronomical files of different formats and types of metadata. All such metadata was parsed and structured, which given us an opportunity to perform the proper metadata mining. Such proper metadata was used in the further research and calculations, where the measurements of each known objects are used to clarify the typical form [38] of image, its orbits, motion parameters and other important astronomical properties on the long historical period.

Also, the *Telescope* software was successfully implemented and installed on the astronomical image processing pipelines in such observatories.

5. Acknowledgements

The authors thank all observatories that provided access to their archives with astronomical data to conduct the current research and test the developed *Telescope* software.

6. References

- [1] H. F. Zhang, C. Chen, J. Wang, et al., Scientific CCD camera for the CSTAR2 telescope in Antarctica, *Journal of Astronomical Telescopes, Instruments, and Systems*, 5 3 (2019). doi: 10.1117/1.JATIS.5.3.036002.
- [2] I. B. Vavilova, et al., Surveys, catalogues, databases, and archives of astronomical data, *Knowledge Discovery in Big Data from Astronomy and Earth Observation*, *Astrogeoinformatics: Elsevier*, (2020) 57-102. doi: 10.1016/B978-0-12-819154-5.00015-1.

- [3] W. D. Pence, L. Chiappetti, C. G. Page, R. A. Shaw, E. Stobie, Definition of the flexible image transport system (fits), version 3.0, *Astronomy & Astrophysics*, 524 42 (2010). doi: 10.1051/0004-6361/201015362.
- [4] J. Furner, Definitions of “metadata”: A brief survey of international standards, *Journal of the Association for Information Science and Technology*, 71 6 (2020). doi: 10.1002/asi.24295.
- [5] R. Klette, *Concise computer vision*, Springer, London, 2014.
- [6] V. Akhmetov, et al., New approach for pixelization of big astronomical data for machine vision purpose, in: *IEEE International Symposium on Industrial Electronics*, 2019, pp. 1706–1710. doi: 10.1109/ISIE.2019.8781270.
- [7] C. Steger, M. Ulrich, C. Wiedemann, *Machine vision algorithms and applications*, John Wiley & Sons, 2018.
- [8] Š. Parimucha, et al., CoLiTecVS – A new tool for an automated reduction of photometric observations, *Contributions of the Astronomical Observatory Skalnaté Pleso*, 49 2 (2019) 151-153.
- [9] L. Mykhailova, et al., Method of maximum likelihood estimation of compact group objects location on CCD-frame, *Eastern-European Journal of Enterprise Technologies*, 5 4 (2014) 16-22. doi:10.15587/1729-4061.2014.28028.
- [10] S. Khlamov, et al., Development of computational method for detection of the object's near-zero apparent motion on the series of CCD-frames, *Eastern-European Journal of Enterprise Technologies*, 2 80 (2016) 41-48. doi: 10.15587/1729-4061.2016.65999.
- [11] V. Akhmetov, et al., Astrometric reduction of the wide-field images, *Advances in Intelligent Systems and Computing*, 1080 (2020) 896–909. doi: 10.1007/978-3-030-33695-0_58.
- [12] V. Savanevych, et al., CoLiTecVS software for the automated reduction of photometric observations in CCD-frames, *Astronomy and Computing*, 40 15 (2022). doi: 10.1016/j.ascom.2022.100605.
- [13] S. Khlamov, V. Vlasenko, V. Savanevych, O. Briukhovetskyi, T. Trunova, V. Chelombitko, I. Tabakova, Development of computational method for matched filtration with analytic profile of the blurred digital image, *Eastern-European Journal of Enterprise Technologies*, 5 119 (2022) 24–32. doi: 10.15587/1729-4061.2022.265309.
- [14] V. Savanevych, et al., Selection of the reference stars for astrometric reduction of CCD-frames, *Advances in Intelligent Systems and Computing*, 1080 (2020) 881–895. doi: 10.1007/978-3-030-33695-0_57.
- [15] R. Gonzalez, and R. Woods, *Digital image processing*, 4th edition, NY: Pearson, 2018.
- [16] L. Kirichenko, A.S.A. Alghawli, T. Radivilova, Generalized approach to analysis of multifractal properties from short time series, *International Journal of Advanced Computer Science and Applications*, 11 5 (2020) 183–198. doi: 10.14569/IJACSA.2020.0110527.
- [17] M. Dadkhah, et al., Methodology of wavelet analysis in research of dynamics of phishing attacks, *International Journal of Advanced Intelligence Paradigms*, 12 (2019) 220-238. doi: 10.1504/IJAIP.2019.098561.
- [18] A. Tantsiura, et al., Evaluation of the potential accuracy of correlation extreme navigation systems of low-altitude mobile robots, *International Journal of Advanced Trends in Computer Science and Engineering*, 8 5 (2019) 2161–2166. doi: 10.30534/ijatcse/2019/47852019.
- [19] N. Yeromina, V. Tarshyn, S. Petrov, et al., Method of reference image selection to provide high-speed aircraft navigation under conditions of rapid change of flight trajectory, *International Journal of Advanced Technology and Engineering Exploration*, 8 85 (2021) 1621–1638. doi: 10.19101/IJATEE.2021.874814.
- [20] M. Ivanov, et al., Effective informational entropy reduction in multi-robot systems based on real-time TVS, in: *IEEE International Symposium on Industrial Electronics*, 2019, pp. 1162–1167. doi: 10.1109/ISIE.2019.8781209.
- [21] V. Akhmetov, et al., Cloud computing analysis of Indian ASAT test on March 27, 2019, in: *Proceedings of the 2019 IEEE International Scientific-Practical Conference: Problems of Infocommunications Science and Technology, PIC S and T 2019*, 2019, pp. 315–318. doi: 10.1109/PICST47496.2019.9061243.
- [22] S. Khlamov, V. Savanevych, Big astronomical datasets and discovery of new celestial bodies in the Solar System in automated mode by the CoLiTec software, *Knowledge Discovery in Big*

- Data from Astronomy and Earth Observation, *Astrogeoinformatics*: Elsevier, 2020, pp. 331–345. doi: 10.1016/B978-0-12-819154-5.00030-8.
- [23] G. K. Adam, et al., Embedded microcontroller with a CCD camera as a digital lighting control system, *Electronics*, 8 1 (2019) 33. doi: 10.3390/electronics8010033.
- [24] J. Andersen, D. Baneke, C. Madsen, *The International Astronomical Union: Uniting the Community for 100 Years*, Springer, 2019.
- [25] H. T. Elshoush, M. M. Mahmoud, A. Abdelrahman, A new high capacity and secure image realization steganography based on ASCII code matching, *Multimedia Tools and Applications*, 81 4 (2022) 5191-5237. doi: 10.1007/s11042-021-11741-y.
- [26] S. Khlamov, et al., Data Mining of the Astronomical Images by the CoLiTec Software, *CEUR Workshop Proceedings*, vol. 3171, 2022, pp. 1043–1055,.
- [27] Y. Zhang, Y. Zhao, C. Cui, Data mining and knowledge discovery in database of astronomy, *Progress in Astronomy*, 20 4 (2002) 312-323.
- [28] S. Bisser, Introduction to the Microsoft Bot Framework, *Microsoft Conversational AI Platform for Developers*, 2021, pp. 25-66. doi: 10.1007/978-1-4842-6837-7_2.
- [29] J. Eckstein, B. R. Schultz, *Introductory relational database design for business, with Microsoft Access*, John Wiley & Sons, 2018.
- [30] V. Savanevych, et al., Mathematical methods for an accurate navigation of the robotic telescopes, *Mathematics*, 11 10 (2023) 2246. doi: 10.3390/math11102246.
- [31] V. Shvedun, S. Khlamov, Statistical modelling for determination of perspective number of advertising legislation violations, *Actual Problems of Economics*, 184 10 (2016) 389-396.
- [32] S. Kashuba, et al., The Simeiz plate collection of the ODESSA astronomical observatory, in: *Proceedings of the 11th Bulgarian-Serbian Astronomical Conference*, 2018, pp. 207-216.
- [33] V. Troianskyi, P. Kankiewicz, D. Oszkiewicz, Dynamical evolution of basaltic asteroids outside the Vesta family in the inner main belt, *Astronomy and Astrophysics*, 672 (2023). doi: 10.1051/0004-6361/202245678.
- [34] I. B. Vavilova, et al., Astrometry and photometry of asteroids from the UkrVO database of astroplates, in: *Proceedings of the IAU*, 15 364 (2021) 239-245. doi: 10.1017/S1743921322000047.
- [35] L. Kirichenko, P. Zinchenko, T. Radivilova, Classification of time realizations using machine learning recognition of recurrence plots, *Advances in Intelligent Systems and Computing*, 1246 (2021) 687–696. doi: 10.1007/978-3-030-54215-3_44.
- [36] V. Baranova, et al., Wavelet Coherence as a Tool for Studying of Economic Dynamics in Infocommunication Systems, in: *IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T)*, 2019, pp. 336-340. doi: 10.1109/PICST47496.2019.9061301.
- [37] I. Kudzej, et al., CoLiTecVS – A new tool for the automated reduction of photometric observations, *Astronomische Nachrichten*, 340 3 (2019) 68-70. doi: 10.1002/asna.201913562.
- [38] V. Savanevych, et al., Formation of a typical form of an object image in a series of digital frames, *Eastern-European Journal of Enterprise Technologies*, 6 120 (2022) 51–59. doi: 10.15587/1729-4061.2022.266988.