

# ESTUDIO MORFOLÓGICO DE REGIONES DE FORMACIÓN ESTELAR USANDO TÉCNICAS DE PROCESAMIENTO DIGITAL DE IMÁGENES ASTRONÓMICAS

Aguirre Manzano, Lizeth (1), Rodríguez Rico, Carlos Alanías (2)

1 Nivel Medio Superior, Escuela de Nivel Medio Superior de Guanajuato | lizgarp91@gmail.com

2 Departamento de Astronomía, División de Ciencias Naturales y Exactas, Campus Guanajuato, Universidad de Guanajuato | ca.rodriguez@ugtomx.onmicrosoft.com

## Resumen

El presente proyecto está orientado hacia el Procesamiento Digital de Imágenes (PDI) en formato FITS, pero desde una perspectiva un poco diferente, ya que no solo se trabajó con la implementación de algunas técnicas de procesamiento digital, también se realizaron algunas rutinas básicas en PHP para que el proyecto fuera un poco más dinámico. Pero para su elaboración primero se realizaron algunas pruebas en el lenguaje de programación R, donde se implementaron algunos scripts cuya función fue "leer" las imágenes y transformarlas a matrices de datos para una fácil manipulación y extracción de información. Posteriormente se les aplicaron algunos algoritmos para poder mejorar algunos aspectos de las imágenes o simplemente hacer más evidentes algunos detalles de estas. La siguiente parte del proyecto consistió en aprender un poco del lenguaje de programación PHP, para poder implementarlo mediante la lectura de scripts que contenían instrucciones precisas para leer la imagen FITS y arrojar una serie de datos como el RGB de cada uno de los píxeles de la imagen, su morfología, etc., además de una cabecera con información básica de la imagen, todo esto para finalmente obtener una imagen en formato PNG y acceder a la información utilizando PHP.

## Abstract

Results of Digital Image Processing (PDI) techniques applied to images in FITS format are presented. The PDI is implemented from a different perspective, since the results of the implementation of these techniques are shown on a web page using some basic routines in PHP to make the project a little more dynamic. As a first step, the images were imported using scripts developed in the R programming language, their function is to "read" the images and transform them into data matrices for easy manipulation and extraction of the information. Subsequently, some algorithms were applied to reduce the noise of the images and apply logarithmic transfer functions to make some details of these more evident. The implementation in PHP consisted of reading the matrices exported from R and corresponding to the FITS image. A matrix containing the information of the RGB image of each of the pixels was generated. Finally, the image was generated in PNG and JPEG format using the corresponding header.

## Palabras Clave

Procesamiento digital de imágenes; FITS; Scripts; PHP; RGB

## INTRODUCCIÓN

El trabajo desarrollado corresponde a la aplicación de técnicas de procesamiento digital de imágenes (PDI) usando R y en el desarrollo de rutinas básicas en PHP para el procesamiento y visualización de las imágenes. Para entender esto, primero debemos de saber que es el PDI y en que consiste el código PHP; pero antes debemos de entender a la imagen no solo como una representación visual de un objeto real, sino como una función bidimensional construida a partir de una matriz binaria  $f(x,y)$  donde  $x$  e  $y$  son las coordenadas espaciales, y  $f(x,y)$  representa la intensidad o nivel de gris/color de la imagen.

En nuestro caso las imágenes procesadas tienen extensión FITS (Flexible Image Transport System), el cual es el formato más utilizado en la Astronomía. FITS es a menudo utilizado para almacenar también datos que no son imágenes, como espectros electromagnéticos, listas de fotones, cubos de datos entre otros. Un fichero FITS podría contener varias tablas anexas, y cada una de ellas podría contener datos de un objeto. Por ejemplo, es posible almacenar imágenes de rayos X y también imágenes pertenecientes al infrarrojo en el mismo archivo FITS.

### Procesamiento Digital de Imágenes (PDI)

La historia del PDI está relacionada con el desarrollo y la evolución de las computadoras, dado que este requiere un alto nivel computacional para procesar y almacenar imágenes. El interés en el Procesamiento Digital de Imágenes se basa esencialmente en dos aspectos: en mejorar la información contenida en una imagen para la interpretación humana y en el tratamiento de los datos en una escena para la percepción autónoma en una máquina. [1]

El PDI hizo su aparición en el mundo científico en la primera mitad del siglo XX. Que con técnicas digitales basadas en algoritmos matemáticos, se logra el mejoramiento para fines de percepción visual humana de imágenes que; en primera instancia no logran ofrecer toda la información visual que poseen. Esta técnica tiene como objetivo mejorar la calidad de las imágenes, hacer más evidentes en ellas ciertos detalles que se desean hacer notar, facilitar la búsqueda o extracción de información de estas imágenes bidimensionales.

Uno de los primeros lugares donde se empezó a realizar el procesamiento digital fue en el Jet Propulsion Laboratory, en 1959, con el propósito de mejorar las imágenes enviadas por los cohetes. [2]

En el Procesamiento Digital de Imágenes se implementan una serie algoritmos con el fin de mejorar ciertos aspectos de la imagen.

#### *Las matemáticas como herramienta fundamental en el PDI*

Las matemáticas nunca dejan de estar presentes en nuestra vida, día a día les encontramos nuevas aplicaciones que anteriormente eran impensables. La aplicación de modelos matemáticos al tratamiento digital de las imágenes ha permitido grandes avances en muchísimos campos del conocimiento, desde la compresión y edición de imágenes hasta la identificación de patrones de enfermedades como el cáncer. [3]. El desarrollo de algoritmos para el procesamiento de imágenes ha cobrado mayor impulso en el ámbito científico, en especial en el campo de la inteligencia artificial para aplicaciones relacionadas con visión por computadora. [4]. Los algoritmos de procesamiento buscan extraer características al convertir una imagen de tamaño fijo en un vector. Esta transformación sirve para detectar patrones. [5]

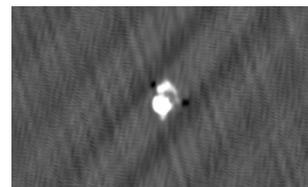
Algunos de los algoritmos que se implementaron en este proyecto fueron principalmente dos, el algoritmo de convolución y el FFT (Fast Fourier Transform). El primero consiste en un tratamiento que se realiza entre la imagen y una matriz llamada "kernel", donde se analiza cada pixel de la imagen y se les asigna un valor diferente, llegando así a una tercera matriz con valores distintos a los iniciales, ayudando de esta manera a la eliminación de ruido (variación de brillo o color en una imagen digital) y la detección de bordes. Mientras que el segundo nos permite calcular la transformada de Fourier discreta; identifica las componentes frecuenciales de la señal, que el caso de la imagen corresponde a los niveles de gris.

## Leyendo imágenes en PHP

**PHP** (Hypertext Preprocessor), es un lenguaje de programación de propósito general de código del lado del servidor originalmente diseñado para el desarrollo web de contenido dinámico, que generalmente se encuentra incrustado dentro del código HTML. Este lenguaje de programación procesa un script solicitado que genera el contenido de manera dinámica, que es interpretado por un servidor web con un módulo de procesador de PHP que genera el HTML resultante. De esta forma nosotros implementamos un script que contenga las instrucciones necesarias para leer una imagen, de tal manera que nos genere un HTML con la imagen en sí o alguna información básica de esta. Aunque el desarrollo de PHP está centrado en la programación de scripts del lado del servidor, se puede utilizar para muchas otras cosas.

## MATERIALES Y MÉTODOS

Para la elaboración de este proyecto fue necesario el uso de softwares que nos permitieran trabajar con imágenes FITS, tales como DS9 y R; el primero nos ayudó a la visualización de la imagen como tal, además de mostrarnos la cabecera de esta donde está la información básica de la imagen, mientras que el segundo nos ayudó a “leer” la matriz bidimensional que corresponde a la imagen. En la imagen 1 se muestra la imagen a procesar con R.



El proyecto se dividió en dos partes: la primera fue un desarrollo de scripts en R y la segunda fue la parte “dinámica del proyecto” hecha con bases del lenguaje PHP. Para la “primera fase del proyecto” fue necesario trabajar en la manipulación de las imágenes FITS en R para poder modificarla de cierta forma que se nos facilitara más su manejo, para ello se realizaron algunos scripts que cumplieran con la función de leer la matriz de datos mediante una serie de funciones ya establecidas. La estructura de nuestros scripts consiste de la siguiente forma: 1) Leer la imagen con extensión FITS y almacenarla en una variable de tipo “array” n-dimensional (e.g. “imagen”); 2) Leer las coordenadas de Ascensión Recta (RA por sus siglas en inglés) y Declinación, transformándolos a vectores; 3) Extraer la matriz de la imagen y almacenarla en otra variable bidimensional que llamamos “imgX”; 4) Aplicar el algoritmo de la transformada rápida de Fourier (FFT) a nuestra imgX; 5) Definir la función de transferencia logarítmica para la imagen; 6) Visualizar la matriz bidimensional (imgX); 7) Aplicar el algoritmo de convolución utilizando una función gaussiana a nuestra imagen para suavizarla y finalmente visualizarla al igual que la imgX.

Al final de la ejecución de los scripts que realizaban estas tareas se comparaba la imgX real con la imgX modificada, todo esto con el fin de comprobar si los cambios aplicados realmente fueron efectivos. En resumen ésta fase solo fue para transformarla en una matriz de dimensiones  $[m \times n]$  donde cada elemento de esta matriz corresponde a un pixel y mediante una serie de operaciones matemáticas mejorar la visualización de la imagen. Las matrices de datos obtenidos a partir de las imágenes FITS, procesadas en R, son exportadas como archivos CSV para un procesamiento posterior.

Para la segunda fase del proyecto (la fase dinámica) fue necesario implementar scripts en el lenguaje de programación PHP y HTML, por lo que fue necesario aprender un poco sobre estos lenguajes de programación; con la implementación de estos dos códigos se desarrollaron los scripts que fueron cargados al servidor por medio del protocolo de transferencia de archivos SFTP. Después del desarrollo de los scripts básicos de PHP se implementaron algoritmos para el PDI en PHP: se leen tres imágenes obtenidas en diferentes longitudes de onda importando los datos CSV previamente exportados desde R. Combinamos la información de las diferentes imágenes para generar un arreglo de dimensión 2X3 y generar los colores correspondientes a cada pixel en escala RGB (Red Green Blue, modelo cromático que consiste en representar distintos colores a partir de la

mezcla de estos tres colores primarios); al tener de cierta forma los colores de la imagen solo faltaba leerlos para que en vez de obtener un arreglo de datos esta vez el resultado final fuera una imagen PNG.

## RESULTADOS Y DISCUSIÓN

Al finalizar el desarrollo de los scripts en R y PHP que tienen como finalidad leer y procesar las imágenes FITS y generar datos CSV e imágenes PNG, decidimos procesar tres imágenes en diferentes longitudes de onda obtenidas por el radiotelescopio VLA (Very Large Array) hacia el objeto G43; que corresponde a una región de gas ionizado por una estrella de tipo O, caracterizadas por tener altas temperaturas, emitir grandes cantidades de radiación ionizante y ser extremadamente masivas; las longitudes de onda de las observaciones procesadas para G43 son 2 cm, 4 cm, y 7 mm. En la Figura 2, 3 y 4 se muestran las imágenes en cada una de estas longitudes de onda, respectivamente.

Para procesar las imágenes fue necesario usar R en donde se creó un script con la matriz de cada una de las imágenes, el otro script fue desarrollado en PHP y sus funciones son: asignarle colores RGB a cada pixel de la imagen combinando los datos de cada imagen FITS (procesada previamente en R) y finalmente se crea la imagen en formato PNG. La generación de la matriz RGB es un proceso que permite asignarle valores a la banda R con base a la imagen de G43 a  $\lambda=2$  cm, colores a la banda G usando la imagen de G43 a  $\lambda=4$  cm y los correspondientes la banda B utilizando los datos de la misma fuente pero a  $\lambda=7$  mm. La superposición nos permite visualizar simultáneamente las diferentes componentes del gas ionizado que rodea a nuestra estrella masiva. Modificamos la función de transferencia de la matriz RGB de la imagen combinada para mejorar la visualización de la región de formación estelar G43.

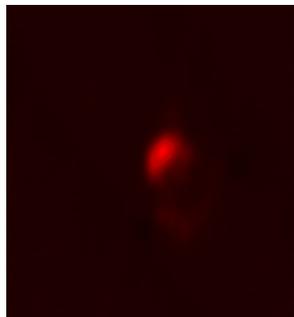
**Tabla 1: Parámetros de las imágenes FITS utilizadas en la generación de imagen RGB.**

Imagen	Escala de placa ("/pixel)	Resolución Angular	Tamaño de imagen
2CMG43.FITS	0.15	2.86"x1.85"	1024x1024
4CMG43.FITS	0.3	1.58"x0.95"	1024x1024
7MMG43.FITS	0.1	1.46"x0.81"	512x512

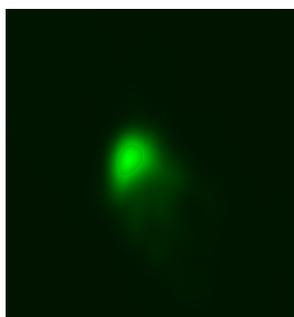
Las imágenes obtenidas en cada una de las diferentes longitudes de onda tienen diferentes características (resolución angular, dimensiones y escala de placa; ver Tabla 1) por lo que se tuvo que realizar un procesamiento previo a la combinación de las mismas. Utilizando la tarea HGEOM de AIPS (Astronomical Image Processing System, sirve para la calibración, edición, formación, mejora y análisis de imágenes y de otros datos astronómicos; además de la reducción de datos de observaciones astronómicas), se realizó la reproyección de los datos y de esta manera lograr consistencia entre tres las imágenes de G43. El tamaño final de cada imagen es de 512x512 pixeles y escala de placa de 0.1"/pixel. En la cual se tomó la imagen con los pixeles más pequeños (y mejor resolución angular) y se ajustaron las otras para que los pixeles coincidieran y las dimensiones también; las imágenes finales utilizadas en la superposición en RGB son de dimensiones 512x512.

Así que en resumen los resultados fueron: el procesamiento de imágenes FITS en diferentes longitudes de onda usando R; exportar las imágenes en formato de texto CSV con la finalidad de acceder a la información usando PHP; ajuste de dimensiones con HGEOM y obtención de la imagen combinada RGB en formato PNG. En la Figura 5 se muestra la imagen resultante RGB tras la combinación de los datos que corresponden a las imágenes de las Figuras 2, 3, y 4. Este desarrollo se realizó con la finalidad de facilitar la visualización de datos astronómicos obtenidos con grandes telescopios (como el radio interferómetro VLA). De esta manera los usuarios

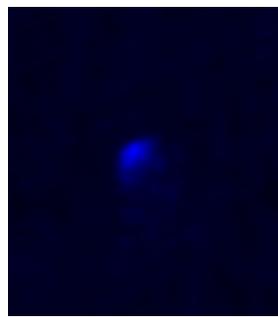
podrán realizar comparación de observaciones realizadas en diferentes longitudes de onda a través de la página del Departamento de Astronomía de la UG.



**IMAGEN 2:** Imagen procesada PNG de G43 en un RGB (255, 0, 0), con  $\lambda= 2$  cm



**IMAGEN 3:** Imagen procesada PNG de G43 en un RGB (0, 255, 0), con  $\lambda= 4$  cm



**IMAGEN 4:** Imagen procesada PNG de G43 en un RGB (0, 0, 255), con  $\lambda= 7$  mm



**IMAGEN 5:** Imagen PNG de las tres longitudes de onda superpuestas de G43.

## CONCLUSIONES

La implementación del PDI en áreas como la Astronomía facilita la visualización y análisis de observaciones realizadas hacia objetos celestes. En este trabajo se presentaron los resultados obtenidos de la implementación de técnicas de PDI en R, PHP y AIPS. Se desarrolló un algoritmo de PDI para combinar imágenes multi-frecuencia de manera consistente: 1) reproyección de imágenes, 2) procesamiento de imágenes FITS para reducción del ruido aplicando suavizado gaussiano y 3) generación de imágenes RGB con visualización en páginas web usando código PHP.

## AGRADECIMIENTOS

Quiero agradecer a la Universidad de Guanajuato y a la Dirección de Apoyo a la Investigación y al Posgrado por brindarme la oportunidad de participar en este Verano de Investigación; a la Doctora Claudia Erika Hernández Morales por el apoyo con los trámites a realizar para poder realizar esta estancia; al Doctor Carlos Alanías Rodríguez Rico por darme la oportunidad de formar parte de este proyecto; a todos mis compañeros de verano que tuve la fortuna de conocer en esta estadía. También quiero agradecer a mis padres y familia en general por su apoyo incondicional y la confianza que tienen en mí.

## REFERENCIAS

- [1] González, R.C. & Wood, R.E. (2018). The Origins of Digital Image Processing. En G & W (Ed), Digital Image Processing: (pp.3-6). New York: Pearson. copyright.
- [2] Sin autor. (Agosto 9, 2012). VI. PROCESAMIENTO DE IMAGENES. Julio 16, 2018, de Biblioteca Digital ILCE Sitio web: [http://bibliotecadigital.ilce.edu.mx/sites/ciencia/volumen2/ciencia3/084/htm/sec\\_9.htm](http://bibliotecadigital.ilce.edu.mx/sites/ciencia/volumen2/ciencia3/084/htm/sec_9.htm)
- [3] Sin autor. (2015). Las matemáticas en el procesamiento digital de imágenes (histórico). Julio 16, 2018, de Konrad Lorenz, Fundación Universitaria, Centro de Investigaciones de Matemáticas e Ingenierías Sitio web: <http://www.konradlorenz.edu.co/es/investigaciones/investigaciones-matematicas-ingenierias/254-las-matematicas-en-el-procesamiento-digital-de-imagenes.html>
- [4] & [5] José Antonio Taquíá Gutiérrez. (2017). El procesamiento de imágenes y su potencial aplicación en empresas con estrategia digital. Interfaces, 10, 11-29.