

*Desarrollo de una plataforma astroinformática para la administración y
análisis inteligente de datos a gran escala*

Apilamiento inteligente de imágenes radio

Mauricio Solar, Marcelo Mendoza, Jorge Ibsen, Lars Nyman, Eduardo Vera, Diego Mardones, Guillermo Cabrera, Paola Arellano, Karim Pichara, Nelson Padilla, Ricardo Contreras, Neil Nagar, Victor Parada, Rodrigo Jara.

Santiago, 16 de agosto de 2013

Resumen

El documento presenta la manera en que se desarrollará el apilamiento con imágenes de radio obtenidas por ALMA. Consta de antecedentes generales del tema, su justificación y ventajas. El fin es desarrollar una herramienta que permita a los astrónomos aprovechar la información existente y analizar aquellas imágenes que individualmente no tienen mayor riqueza pero que al apilarlas se logran resultados sorprendentes.

Palabras claves: FITS, apilamiento de imágenes, interferometría.

1. Resumen ejecutivo

Atacama Large Milimeter/Submillimeter Array (ALMA), es el mayor proyecto astronómico que existe, produce una fracción creciente de datos por día que va siendo almacenada en repositorios que después de cierto tiempo está disponible para cualquier investigador. Aquellos datos corresponden a resultados de una investigación específica, información que puede ser reutilizada para otros fines de interés. Una forma de aprovecharla es realizando el apilamiento de imágenes, técnica conocida hace mas de dos décadas, donde su primera aplicación se conoció con Caillault y Helfand en el estudio de emisiones de rayos X, que les ayudo a determinar la edad de la estrella respecto al deterioro de la emisión. El apilamiento consiste en generar una imagen que represente a un conjunto de imágenes previamente seleccionadas. Esto contrarresta los efectos generados por el instrumento o las condiciones atmosféricas, disminuyendo considerablemente el ruido en la imagen y aumentando su profundidad.

Tabla de contenidos

1. RESUMEN EJECUTIVO	2
2. METODOLOGÍA DE TRABAJO	5
3. OBJETIVOS DEL PROYECTO.....	5
3.1 OBJETIVO GENERAL.....	5
3.2 OBJETIVO ESPECÍFICO	5
4. DESCRIPCIÓN DEL PROBLEMA.....	6
4.1 MOTIVACIÓN.....	6
4.2 ENUNCIADO DEL PROBLEMA	7
4.3 ESTADO DEL ARTE	8
5. DESCRIPCIÓN DE LA SOLUCIÓN PROPUESTA	9
5.1 CARACTERÍSTICAS DE LA SOLUCIÓN	9
5.2 PROPÓSITOS DE LA SOLUCIÓN.....	9
5.3 ALCANCES O LIMITACIONES	10
6. METODOLOGÍAS, HERRAMIENTAS Y AMBIENTE DE DESARROLLO	10
6.1 METODOLOGÍA A UTILIZAR.....	10
6.2 HERRAMIENTA DE DESARROLLO.....	10
6.3 AMBIENTE DE DESARROLLO	11
7. PLAN DE TRABAJO	11
BIBLIOGRAFÍA	12

Índice de Figuras

1. PROCESO DE APILAMIENTO DE IMÁGENES	8
2. CARTA GANTT DEL PROYECTO	11

2. Metodología de trabajo

El presente documento se realizó a partir de la reunión realizada con el astrónomo Neil Nagar en la Universidad de Concepción, y de las reuniones no presenciales previas. El estudio personal y el apoyo constante del equipo de trabajo detrás de Neil, ha sido fundamental para esclarecer y definir el trabajo a desarrollar.

El desarrollo general de este documento se realizó en un periodo muy acotado, comenzó con el estudio guiado del tema, continuando con la reunión presencial con Neil, la cual fue el punto clave para hacer fluir todas las ideas. Posteriormente he recibido ayuda por parte del equipo de trabajo de Neil, quienes me han seguido orientando y me han dado las facilidades para desarrollarme en esta área.

3. Objetivos del proyecto

3.1 Objetivo general

Desarrollar una aplicación que permita el apilamiento inteligente de imágenes radio-interferométricas generadas por el observatorio ALMA[1] para el análisis de objetos con características similares para su posterior estudio.

3.2 Objetivo específico

- a. Diseñar un algoritmo que produzca apilamiento de imágenes.
- b. Seleccionar arbitrariamente un conjunto de imágenes para experimentación a partir de observaciones captadas por el telescopio espacial Herschel[2].
- c. Identificar e implementar parámetros de selección asociados a la calidad, resolución y espacio/tiempo de una imagen.
- d. Recortar en cada imagen el objeto de interés para focalizar el apilamiento en determinado objeto.
- e. Diseñar experimento computacional para evaluar el desempeño del algoritmo con imágenes ya estudiadas.
- f. Generalizar la aplicación de este algoritmo a las imágenes obtenidas por ALMA.

4. Descripción del problema

4.1 Motivación

Las imágenes astronómicas generadas por cualquier telescopio, se ven limitadas por varios factores. Entre ellos, las condiciones atmosféricas, la resolución del telescopio o la razón señal/ruido[3] en cada pixel de la imagen. La señal captada en cada pixel, depende del flujo de la fuente astronómica y el tiempo de integración. El ruido existente en cada pixel depende netamente del instrumento de observación. Por ejemplo, en un circuito de carga acoplada (CCD)[4], el ruido se genera por la agitación térmica de los portadores de carga (conductores), por la lectura del CCD y el "fondo" de emisión del cielo y/o telescopio. En general, un mayor tiempo de integración en una fuente óptica proporciona imágenes con una señal mas expuesta al ruido.

En la obtención de imágenes radio-interferométricas, el ruido se genera principalmente por la agitación térmica de los receptores ubicados en cada telescopio. Aquí, interacciones mas largas dan lugar a imágenes con mejores niveles de señal respecto al ruido, debido a que la integración ya produce una disminución de este. Para un telescopio como ALMA, la relación señal/ruido es inversamente proporcional a la raíz cuadrada del producto de la anchura de banda y el tiempo de integración total. Es decir, para un ancho de banda fijo, la duplicación de la señal al ruido requiere cuatro veces el tiempo de integración.

En el caso del observatorio ALMA, en muchos casos, no se detectan objetos astronómicos individuales en una imagen, es decir, la señal de este objetivo es más pequeño que el ruido (o tres veces el ruido) en la imagen. Sin embargo, si uno tiene una muestra de los objetivos (por ejemplo, 1.000 fuentes) con posiciones conocidas con precisión, todos los cuales estan en la imagen, se podría extraer subimágenes centradas en cada posición y la media de todas las 1000 sub-imágenes. La imagen promedio es equivalente a una imagen con 1000 veces el tiempo de integración, es decir, el ruido en la imagen se reduce por un factor de la raíz cuadrada de 1000 o aproximadamente 30. Por tanto, es más probable que la imagen promediada mostrara una detección de señales muy debiles que requeririan de dias o meses de observacion (la señal en la imagen promediada es la señal media de las 1000 fuentes, mientras que el ruido es 30 veces menor que el

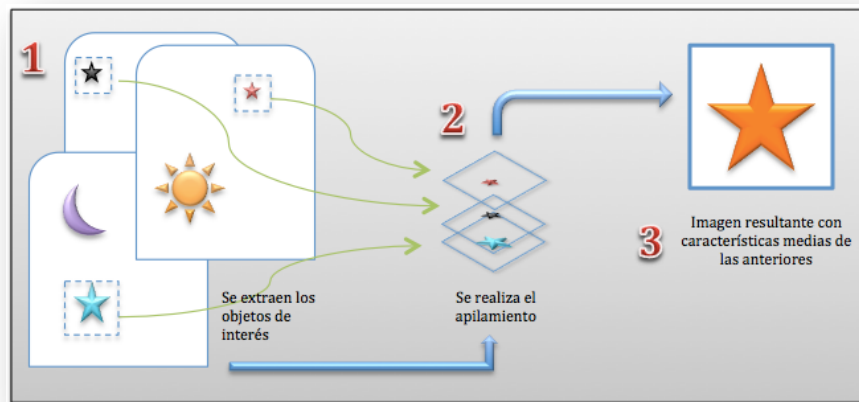
ruido en una cualquiera de las imágenes de las 1000 fuentes). Esta técnica, conocida como apilamiento, es una herramienta muy potente y se utiliza, por ejemplo, en imágenes de rayos X, formación de imágenes ópticas e imágenes radio. El apilamiento da una idea de las características medias de una muestra, algo muy útil para el análisis.

Como el archivo de ALMA sigue creciendo, una fracción cada vez mayor del cielo será capturada. Con una muestra suficientemente amplia de objetos astronómicos, por ejemplo, un millón de cuásares con posiciones precisas del Sloan Digital Sky Survey, SDSS[5], se encuentra que muchos objetos de la muestra estarán en áreas que ya han sido observadas por ALMA. Incluso si no se detectan de forma individual, se puede apilar todos los objetos que han sido vistos por ALMA y obtener un flujo promedio para esta muestra. Si la muestra también tiene desplazamientos al rojo[6], el apilamiento se puede realizar en el espacio de velocidades con el fin de obtener las propiedades medias de cualquier línea de emisión dada.

4.2 Enunciado del problema

Para abordar apilamiento de imágenes, los astrónomos realizan típicamente de forma manual una serie de procedimientos. En la primera etapa, se buscan diversas fuentes de imágenes, por ejemplo, galaxias elípticas, de las cuales se dejarán solo aquellas que no posean ruido térmico intractable, es decir, fuentes luminosas potentes que contaminan sus vecinas. Una vez que ya tienen las imágenes pasan a una segunda etapa, que consiste en recortar de la imagen general el punto de interés a estudiar, manteniendo las mismas coordenadas del cielo. Ya teniendo los recortes, ‘nivelan’ las imágenes, ajustando su resolución, longitud de onda, ubicación especial, etcetera, bajo un valor común. Una vez realizado todos estos pasos, ya es posible apilar con sentido.

En la Fig.1, se busca graficar de manera general la forma mediante la cual se toma un objeto de interés de cada una de las imágenes pertenecientes a una muestra, todo ello con el fin de apilar dichos objetos y así conseguir mayor profundidad en la imagen resultante.



“Fig 1: Proceso de apilamiento de imágenes”

El problema central de la serie de pasos realizados por el astrónomo está en la manualidad de las operaciones. No solo cada astrónomo lo lleva a cabo con su propio estilo de trabajo, si no que además, el diseño de sus programas no obedece el diseño de algoritmos que típicamente se sigue en el ámbito de la Ingeniería Informática. Un algoritmo como el que se propone desarrollar en esta memoria, permitiría a los astrónomos realizar el apilamiento automático de imágenes, incurriendo en menores costos en la detección de objetos.

4.3 Estado del Arte

Todas las observaciones de imágenes astronómicas poseen un umbral de sensibilidad por debajo del cual los objetos no son detectables. Si uno tiene razones para creer que en esa posición del cielo puede existir algo, basándose en otras observaciones, se puede aplicar apilamiento a una zona delimitada a un cierto rango de longitudes de onda, pudiendo variar dicha longitud, y así conseguir el flujo medio de emisiones.

En una aplicación inicial de esta técnica, Caillault y Helfand, detectaron el flujo medio de rayos X de estrellas G en Pleiades, utilizándola para determinar la edad respecto al deterioro de la emisión[7]. Los detectores de emisiones de rayos X han estado disponible durante más de dos décadas, por lo que el apilamiento se ha convertido en una técnica de análisis estándar. Las aplicaciones comenzaron desde la determinación de la

luminosidad media de rayos X, por ejemplo las galaxias normales[8], galaxias Lyman[9] y las fuentes de radio[10], para determinar la emisión de rayos X de grupos distantes en el Röntgensatellit (ROSAT) All-Sky Survey[11].

Los detectores digitales lineales han llegado a dominar los estudios del cielo óptico e infrarrojo. La técnica de apilamiento ha sido ampliamente adoptada: por ejemplo, Zibetti et al.[12] detectaron la luz intracluster apilando Sloan Digital Sky Estudio 683 (SDSS6); Lin et al.[13], apiló datos sobre las galaxias del cúmulo Two Micron All Sky Survey (2MASS); Hogg et al.[14] apiló datos del Keck IR para conseguir galaxias de colores tenues; Minchin et al.[15] fue tan lejos como para apilar películas digitalizadas desde el telescopio Schmidt del Reino Unido.

5. Descripción de la solución propuesta

5.1 Características de la solución

Existen variados software que actualmente realizan apilamiento de imagen en diferentes escalas, pero en el caso de la astronomía, la serie de pasos previos para lograr aplicar el apilamiento, dificulta la tarea. Por tal motivo, se desarrollará un software capaz de aprovechar las herramientas existentes que permiten procesar previamente las imágenes antes de ser apiladas, y así unificar todas estas funciones como un solo producto.

5.2 Propósitos de la solución

Crear una herramienta que permita a los astrónomos concentrarse en la observación mas que en como tratar la imagen para su posterior análisis. Es decir, desarrollar un software inteligente que permita unificar todas aquellas herramientas que llevan a aplicar el apilamiento. Esto aumentaría considerablemente la producción de los científicos, se aprovecharía potencialmente toda la información existente, y disminuirían los costos asociados a observar.

5.3 Alcances o limitaciones

En cuanto a alcances y limitaciones, se debe tener presente que el software: trabaja unicamente con imágenes cuya extensión de archivos sea ‘FITS’[16]; corresponde a una herramienta de filtro, por lo que no almacena información; trabaja con imagines de tipo radio; no cuenta con la información requerida para el apilamiento, esta debe ser entregada por el operador; por ultimo, el resultado obtenido dependerá netamente de la calidad de los datos de entrada.

6. Metodologías, herramientas y ambiente de desarrollo

6.1 Metodología a utilizar

En este punto se describe la metodología a usar para lograr los objetivos de investigación. Debido a la naturaleza del objetivo de desarrollo, el método mas adecuado es el "Desarrollo incremental". Es decir, el proyecto se dividirá en etapas, en donde cada una de ellas entregará una nueva funcionalidad, es decir, cada nivel consiste en requerimientos, diseño, codificación y pruebas. Esto permitirá atender de forma organizada, oportuna y eficiente los requerimientos, asegurar la calidad del producto resultante, proveer visibilidad del progreso, asegurar los compromisos con el cliente para lograr en última instancia la satisfacción de ellos.

Las funcionalidad previas con las que contará el sistema antes de realizar el apilamiento, son claras, por lo que las metas a corto plazo a definir no son lejanas.

6.2 Herramienta de desarrollo

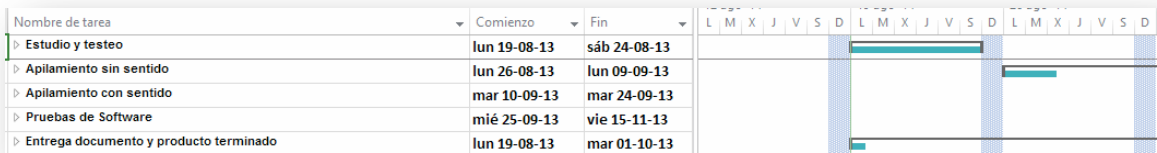
El lenguaje Python está teniendo bastante impulso en ámbitos científicos durante los últimos años, y uno de los campos donde es bastante activo es en astronomía. Pyfits es un modulo de Python el cual proporciona un interfaz a los archivos con formato FITS, el cual se puede utilizar para el análisis y posterior escritura de este tipo de archivos. Por otro lado también existe un software llamado CASA[17], el cual posee como objetivo principal el apoyar las necesidades de post procesamiento de datos de la próxima generación de radio telescopios como ALMA y VLA[18]. CASA utiliza Python, por lo

tanto se puede aprovechar en Python todas las funciones que posee CASA para el tratamiento de imágenes de radio generadas por ALMA.

6.3 Ambiente de desarrollo

El desarrollo de este trabajo se apoya en los recursos humanos dispuestos por la Universidad de Concepción, conformado por un par de astrónomos y alumnos realizando su doctorado en el área de la astronomía, los cuales depositan su confianza en un proyecto de interés común.

7. Plan de trabajo



“Fig 2: Carta Gantt del Proyecto”

Bibliografía

- [1] Heisig, M. (2007, abril). Cerca del cielo: Patrimonio biológico en la zona de alma. [Consulta: 3 de agosto 2013]. <
http://www.almaobservatory.org/images/pdfs/alma_observatory_book.pdf>
- [2] Participación del IGN en el Telescopio Espacial Herschel. (2006, Junio).
[Consulta: 3 de agosto 2013]. <
http://www.fomento.gob.es/NR/rdonlyres/46003C7C-A662-422E-934F-1E4D3FA88248/21214/bol26_p4.pdf>
- [3] Doreste, J. (2010, 7 de junio de). Relación Señal a Ruido en imágenes CCD astronómicas. [Consulta: 3 de agosto 2013]. <http://www.astrosurf.com/cometas-obs/ArtSoftUtil/Articulos/SNR_JLDoreste.pdf>.
- [4] Violat, F. (2004, 21 de Marzo). Astronomía con CCD. Espacial. [Consulta: 3 de agosto de
2013] <http://www.espacial.org/astronomia/tecnologia/astronomia_ccd1.htm>.
- [5] Sloan Digital Sky Surbey. [Consulta: 3 de agosto 2013]. <<http://www.sdss.org>>
- [6] Zorzano, J. (2008, septiembre). Desplazamiento hacia el rojo (Redshift).
<[http://faii.industriales.upm.es/dfaii/Docencia/Material%20Docente/Introducción%20a%20la%20Cosmolog%C3%ADa%20I/II-Ley%20de%20Hubble%20v1.pdf](http://faii.industriales.upm.es/dfaii/Docencia/Material%20Docente/Introduccion%20a%20la%20Cosmolog%C3%ADa%20I/II-Ley%20de%20Hubble%20v1.pdf)>
- [7] Caillault, J.-P., & Helfand, D. J. 1985, ApJ, 289, 279
- [8] Brandt, W. N., et al. 2001b, AJ, 122, 1
- [9] Brandt, W. N., Hornschemeier, A. E., Schneider, D. P., Alexander, D. M., Bauer, F. E., Garmire, G. P., & Vignali, C. 2001a, ApJ, 558, L5
- [10] Georgakakis, A., Hopkins, A. M., Sullivan, M., Afonso, J., Georgantopoulos, I., Mobasher, B., & Cram, L. E. 2003, MNRAS, 345, 939
- [11] Bartelmann, M., & White, S. D. M. 2003, A&A, 407, 845
- [12] Zibetti, S., White, S. D. M., Schneider, D. P., & Brinkmann, J. 2005, MNRAS, 358, 949
- [13] Lin, Y.-T., Mohr, J. J., & Stanford, S. A. 2004, ApJ, 610, 745

- [14] Hogg, D. W., Neugebauer, G., Armus, L., Matthews, K., Pahre, M. A., Soier, B. T., & Weinberger, A. J. 1997, AJ, 113, 474
- [15] Minchin, R. F., et al. 2003, MNRAS, 346, 787
- [16] Morisset C.,L.(2012, 25 de marzo). Play with FITS files.[Consulta: 17 de mayo de 2013].< <http://python-astro.blogspot.com/2012/03/play-with-fits-files.html>>
- [17] Kern, J.(2013, 28 de mayo). CASA Synthesis & Single Dish Reduction Reference Manual & Cookbook. [Consulta: 7 de junio 2013]< <http://casa.nrao.edu>>
- [18] Very Large Telescope.(2012, 6 de Diciembre).[Consulta: 26 de julio de 2013].< <http://www.eso.org/public/teles-instr/vlt.html>>