

D. Cristóbal-Hornillos, A.J. Cenarro, N. Gruel, J.L.Lamadrid, N. Maicas,  
A. Marín-Franch, M. Moles, J. Montero, S. Rueda, K. Viironen  
CENTRO DE ESTUDIOS DE FÍSICA DEL COSMOS DE ARAGÓN

## Datos del Observatorio de Javalambre

Usando el T250 del Observatorio Astrofísico de Javalambre (OAJ) se va a realizar un survey de 8000 grados cuadrados en un sistema de 42 filtros de  $\sim 100-120\text{\AA}$  y 2 filtros anchos. La cámara con la que se va a realizar este survey contiene un mosaico de 14 CCDs para ciencia [Fig.1], cada uno de ellos viendo el cielo a través de un filtro distinto. Los CCDs a usar son de  $9.2k \times 9.2k$  píxeles de 10 micras, lo que da una resolución de  $0.23''/\text{pix}$ . Cada imagen producida por uno de estos CCDs se almacenará en formato FITS ocupando  $\sim 162.0\text{MB}$ .

Considerando el tamaño de píxel el survey total equivale a  $2.0T\text{pix}$  ( $\sim 3.6T\text{Bytes}$  en formato 16bits). Como en cada posición del survey se van a tomar 5 exposiciones individuales, los datos sin procesar en el sistema de 42 filtros ocuparán  $\sim 770\text{TBytes}$ . También hay que tener en cuenta el espacio necesario para albergar los datos de calibración.

Usando el T80 se realizará, durante los 2 primeros años de operación, un survey de 8000 grados cuadrados en 14 filtros intermedios para hacer la calibración del JPAS. El telescopio de 80cm estará equipado con una cámara con un CCD similar a los de la cámara del T250. Con una resolución de  $0.56''/\text{pix}$  los datos originales de este survey ocuparán  $\sim 43\text{TBytes}$ . Un resumen de los datos relacionados con el survey JPAS que van a obtenerse durante los primeros años de operación de los telescopios del OAJ puede verse en la tabla [Tab.1]

Tab.1	T80	T250
Exposición individual	$\sim 162\text{ MB}$	(x14) $\sim 2.2\text{ GB}$
Noche (500 exp)	$\sim 80\text{ GB}$	(x14) $\sim 1.1\text{ TB}$
Survey datos raw (16 bits)	(x 14 filtros, 5 dither) $\sim 43\text{ TB}$	(x 42 filtros, 5 pos.) $\sim 770\text{ TB}$
Survey datos procesados (32 bits + 16bits wm)	(x 14 filtros, 5 dither) $\sim 144\text{ TB}$	(x 42 filtros, 5 pos.) $\sim 2.2\text{ PB}$
FINAL combinadas	$\sim 29\text{ TB}$	$\sim 450\text{ TB}$
Datos de calibración raw ( $\sim 20$ por noche)	(x 200 noches) $\sim 1.3\text{TB}$	(x14, 800 noches) $\sim 35\text{ TB}$
Imágenes de calibración combinadas < 4 imágenes por noche, 32 bits)	(x 400 noches) $\sim 510\text{ GB}$	(x14, 800 noches) $< 14\text{ TB}$
Base de datos con los catálogos de objetos		$7 \times 10^7$ to $i=24$ , $8\text{KB}$ objeto $\rightarrow 40\text{ TB}$

Fig.1



## Unidad de procesado de datos

### Hardware de almacenamiento

Los datos sin procesar obtenidos en los primeros 4 años de operación del OAJ serán de  $\sim 1.0\text{PB}$  bytes. Los requerimientos de hardware para almacenamiento pueden identificarse en la [Tab.1]. En la Unidad de Procesado y Análisis de Datos (UPAD) se dispondrá de la infraestructura necesaria para poder almacenar una copia operativa de los datos sin procesar e imágenes combinadas de calibración (bias, flatfields, patrones del fringing). De momento la solución de copia de seguridad será usando una matriz de discos extraíbles en el observatorio y otra en la sede del CEFC. Los datos de cada noche de observación se descargarán inmediatamente a la UPAD tras un primer análisis de integridad y clasificación en el observatorio. Para la descarga de los datos existirá una conexión mediante antenas unidireccionales. En el observatorio habrá un sistema que permitirá almacenar hasta 1 mes de observaciones.

El almacenamiento de las imágenes individuales reducidas e imágenes intermedias no está previsto salvo en las fases iniciales de verificación del sistema de reducción. En cuanto a las imágenes combinadas, solo está previsto el almacenamiento de la última versión. En cada versión de las reducciones se almacenarán las imágenes de calibración (bias, flat, patrones de fringes), las máscaras, y catálogos, tanto para imágenes individuales como finales. Estimamos que los catálogos y máscaras asociados a las imágenes individuales ocupan  $\sim 1\%$  de lo que ocupan las imágenes individuales, mientras que su obtención supone un costo computacional importante. Durante el procesamiento de las imágenes individuales deben almacenarse todas las operaciones e identificaciones de las versiones del software para permitir recuperar las imágenes procesadas desde los datos originales e imágenes de calibración.

### Hardware de procesamiento

Una vez se han recibido en la UPAD los datos de una noche se lanzará automáticamente la reducción de los mismos. Es un requerimiento que los datos sean procesados durante el tiempo diurno para alimentar las bases de datos y actualizar los apuntados para la noche siguiente. Hemos hecho una estimación del tiempo que se requiere para la corrección de los efectos instrumentales. En el caso de los filtros azules es más restrictiva dado el mayor número de imágenes recogidas:

- Los pasos estándar de reducción requieren al menos 3 operaciones por píxel, en una buena noche de 10 horas con tiempos de exposición  $< 60\text{seg}$  podemos obtener  $\sim 500$  imágenes de ciencia. Una velocidad de acceso a disco  $150\text{MB/s}$  conllevaría  $\sim 2.2$  horas de acceso a disco.
- Generar los catálogos para las imágenes individuales usando Sextractor (Bertin & Arnouts, 1996) ( $\sim 2\text{Mpixel/s}$  en un núcleo de  $2\text{GHz}$ ) llevaría a unas 15 h de procesamiento.
- Generar las máscaras de rayos cósmicos con L.A.Cosmic/IRAF (van Dokkum, 2001), ( $0.1\text{Mpixel/s}$  en  $2\text{GHz}$ )  $\rightarrow 150\text{ h}$ .

En el caso de la cámara del T250 se obtendrán 14 imágenes en filtros distintos procedentes del mosaico de detectores. Debido a que la correcciones a aplicar sobre las imágenes son diferentes para cada CCD, la reducción de cada CCD se hará en una máquina distinta. El sistema de discos de cada máquina de reducción será usado a modo de espacio de trabajo temporal con lo que pueden configurarse en un RAID0. El número de procesadores y RAM de cada máquina se escalará para reducir los tiempos de ejecución anteriores a unas 10 horas.

## Software de procesado de datos

Los objetivos para el software de tratamiento de datos son los siguientes:

- Validar la integridad de los datos y clasificar automáticamente las imágenes del T80 y T250 y proceder con la reducción de los datos.
- Corregir los efectos instrumentales de los CCDs
- Calibrar la astrometría y fotometría de las imágenes individuales.
- Validar las imágenes individuales con los requerimientos necesarios para el survey.
- El sistema, tras el procesamiento de los datos de cada noche, debe actualizar una base de datos con el estado del survey. Esto se usará para identificar los apuntados de los telescopios para las siguientes noches.
- Combinar las imágenes individuales de una zona de cielo determinada y generar el mapa de pesos asociado.
- Extraer los catálogos de objetos.
- Verificar periódicamente la calibración de las imágenes y en caso necesario recalibrar los datos para mejorar la homogeneidad de los puntos cero dentro del survey.

Dado que el sistema debe trabajar de manera automática, y que las imágenes individuales corregidas no van a almacenarse, son requerimientos al sistema:

- Debe producir logs y figuras que serán archivados en bases de datos y presentados en formato web, lo que permitirá una fácil validación del proceso de reducción: calidad fotométrica, astrometría, validación de las imágenes. Así como visualizar el grado de completitud del survey.
- Debe permitir trazar las imágenes individuales y de calibración (bias, darks, flatfields, fringing), los offset y factores de escala que se han formado parte del proceso de reducción para la obtención de una imagen combinada final. Para ello las operaciones sobre las imágenes individuales originales y la direcciones a las imágenes de calibración deben almacenarse en una base de datos.

La figura [Fig. 2] muestra un esquema con las distintas partes del sistema de tratamiento de datos descritas arriba.

