

Digitalización 3D automática con láser escáner, fotogrametría y videogrametría. El caso práctico del Templo de Diana (Mérida).

Pedro Ortiz Coder y Beatriz del Pino Espinosa

Instituto de Arqueología de Mérida-CSIC. Mérida. España

Resumen

La fotogrametría y la videogrametría se presentan como técnicas no intrusivas de documentación gráfica del patrimonio, con una aplicación creciente que se impone a otros métodos documentales como el láser escáner, en cuanto a economía y rapidez. En el artículo describimos el proceso metodológico de digitalización de un monumento complejo, como el caso del Templo de Diana, nombrado Monumento Nacional desde 1932 e inserto dentro del conjunto arqueológico de Mérida, incluido en la lista Patrimonio Mundial desde 1993, que ha perdurado con protagonismo reseñable a lo largo de su historia en la ciudad emeritense. Para ello, hemos empleado las técnicas anteriormente aludidas constituyendo un proceso sencillo, mediante las cuales se obtienen resultados para distintos tipos de aplicaciones y necesidades.

Palabras Clave: DOCUMENTACIÓN DEL PATRIMONIO, FOTOGRAMETRÍA, VIDEOGRAMETRÍA, TEMPLO DE DIANA.

Abstract

Photogrammetry and videogrammetry are presented as non-contact heritage documentation techniques, with an increased application which is being imposed to other documentation methods in order to economy and rapidity. In this paper, we describe the methodological process of digitalization of a complex monument such as the case of the Temple of Diana in Mérida, National Monument since 1932, inserted in the Archaeological Ensemble of Mérida (Spain), and included in the List of World Heritage since 1993, which has remain with an important prominence in the city through its history. For this purpose, we had used some of the techniques before mentioned, becoming a simple process through which we can obtain different results for different kinds of applications and needs.

Key words: HERITAGE DOCUMENTATION, PHOTOGRAMMETRY, VIDEOGRAMMETRY, TEMPLE OF DIANA.

1. Introducción: la documentación gráfica del Templo de Diana de Mérida

El Templo de Diana de Mérida tiene la peculiaridad de haber estado presente en el urbanismo de Mérida desde época romana hasta la actualidad, viéndose sometido a una serie de metamorfosis en su apariencia, en relación a sus diversos usos a lo largo de su historia.

Sin entrar en una descripción detallada del Templo, el cual ha sido objeto de estudio en otras ocasiones (especialmente en ÁLVAREZ y NOGALES, 2003; y AYERBE et alii, 2009), apuntamos que se trata de un templo hexástilo y períptero que presidía el llamado complejo forense de la Colonia de *Augusta Emerita*. Tras el abandono de su uso primigenio, el edificio sería objeto de diversas reutilizaciones que se identifican principalmente como espacio doméstico desde la construcción del Palacio de los Corbos hacia finales del siglo XV, manteniendo dicha finalidad hasta finales del siglo XX. En ese momento se iniciaría una última etapa que significaría su puesta en valor como elemento patrimonial.

El proceso comenzó con la declaración como Monumento Nacional de las Ruinas de Mérida junto a las de Itálica por R.O. en diciembre de 1912, cuyo carácter difuso respecto al Templo objeto de estudio se concretó en el informe redactado el 22 de abril de 1932, donde se reconocía la declaración explícita del

edificio, que conservaría el carácter privado que le había definido hasta finales del siglo XX. En 1972 se inició el proceso de expropiación del solar en el que se encontraba el edificio, dando comienzo una serie de intervenciones dentro del marco de diversos proyectos de investigación y adecuación. Las excavaciones comenzaron con una primera intervención codirigida por José Álvarez Sáenz de Buruaga y José María Álvarez Martínez en 1972, continuada por éste último con otras excavaciones hasta 1987, y Félix Palma desde 2001 a 2011 (ÁLVAREZ y NOGALES, 2003; AYERBE et alii, 2009). Paralelamente se realizaron una serie de restauraciones por parte de José Menéndez-Pidal desde 1972, y Dionisio Hernández Gil desde 1985. Finalmente la adecuación de su entorno, realizada por el arquitecto José María Sánchez García, se inauguró en 2011, pretendiendo evocar en parte la apariencia originaria en su contexto romano, y resaltar la preeminencia del Monumento en el urbanismo emeritense.

La relevancia del edificio hizo que tradicionalmente fuera objeto de captación en forma de diversos formatos gráficos: dibujos, planos, grabados, fotografías, así como retratos literarios, surgiendo en su mayor parte gracias a la iniciativa de la monarquía, desde la primera representación del mismo de la que hay constancia, por parte de Antón Van den Wyngaerde por encargo de Felipe II, a mediados del siglo XVI (*figura 1*). Por otro lado también su plasmación surgirá por iniciativa personal de viajeros, eruditos e investigadores. Este repertorio gráfico, con mayor o menor grado de fidelidad respecto al original,

dejándose llevar en muchos casos por estereotipos, tendencia a lo pintoresco, y reiteración de imágenes, nos permite elaborar una crónica visual de su percepción y evolución, que se ve completada mediante el proceso de digitalización realizado (DEL PINO, e. p.).

De este modo, podemos crear una narración gráfica “imaginando” el edificio desde su construcción como parte del *Forum Coloniae* de *Augusta Emerita*, el cual pasaría por un periodo de abandono hasta su uso como espacio doméstico, continuando con las mencionadas intervenciones realizadas en el edificio y su entorno, para su puesta en valor como monumento, remarcando la relevancia como icono de la ciudad.

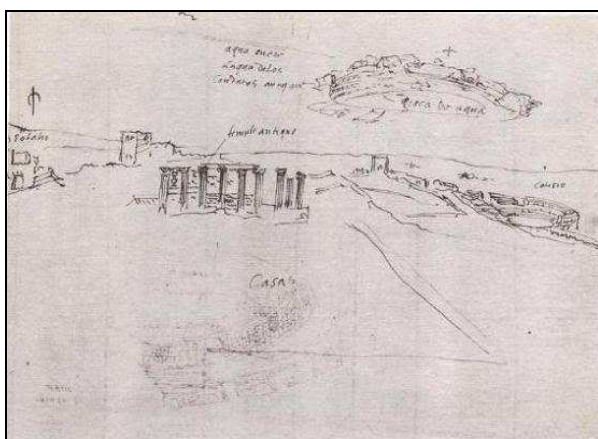


Figura 1. Templo de Diana de Mérida dibujado por A. den Wyngaerde hacia 1567 (Kagan, 1986).

1.1 Diversidad en técnicas de captura métrica

La captura de información métrica es una forma de documentar el Patrimonio, ejemplificado en este artículo en un edificio concreto, pero aplicable a áreas más amplias u objetos de diversas dimensiones. Estas técnicas se realizan a partir de sensores activos o pasivos, constituyendo una forma de inspección y estudio de los detalles de los objetos o elementos documentados. La medición detallada de la superficie puede facilitar un análisis interactivo con el objeto, supone un elemento de consulta y conocimiento que ofrece una herramienta fundamental de establecimiento de hipótesis físicas como la imposición de pesos, estudios de deformaciones y desviaciones, etc, pero también análisis para restauraciones. Por otra parte el modelado 3d supone una réplica digital, y se puede materializar físicamente en segundo término.

La captura de la realidad es una forma de interactuar matemáticamente con ésta. Permite simular condicionamientos muy diversos sobre la realidad que no serían posibles de otra forma, o también planificar obras o casos de gran complejidad, con un menor gasto de recursos.

La captura de objetos 3d también debe clasificarse en métodos de captura masiva de información, donde el sensor capta toda la información posible y la discriminación o es mínima o no existe y, métodos de captura discriminatoria, en los que el usuario elige la medida o los puntos a tomar ralentizando el proceso pero discriminando mucha información posiblemente inútil para el usuario. La técnica que analizamos en este artículo es la de

captura masiva de información, y más concretamente, la que proporciona información totalmente automática.

Es importante definir una nueva clasificación dentro de los métodos de captura masiva a través de imágenes: 1. Métodos con algoritmia basada en la visión computacional (CV), caracterizados por la prioridad de sus procesos en la consecución de un método visualmente atractivo y rápido pero, normalmente deficiente en la precisión (véanse sistemas de modelado basado en web tipo Photofly o directamente con algoritmia CV como puede ser Photoscan). 2. Métodos basados en algoritmia fotogramétrica, caracterizados por utilizar formulaciones más robustas en cuanto a la precisión, donde el coste computacional es mayor aunque los resultados visuales, aun no siendo preferentes, sean cada vez mejores (aplicaciones con este tipos de características son 4e software o Photomodeller scanner).

1.2 Objetivos del proyecto

Con este proyecto tratamos de conseguir una perspectiva metodológica de captura métrica automática y no intrusiva para un caso práctico concreto: El templo de Diana. A través de la contextualización del edificio, tratamos de definir la utilización y adecuación de unas u otras técnicas a partir de una comparación metodológica y de resultados. Sólo a partir de un conocimiento general de las diversas técnicas, se podrá llegar a una elección óptima para cada caso concreto de documentación gráfica.

2. Metodología

El proceso de digitalización empleado tiene como objetivo la documentación gráfica de una forma automática para facilitar y democratizar los procesos de captura métrica a todo tipo de usuarios, dada la disminución de la complejidad. Estos métodos constituyen técnicas no intrusivas que permiten conocer con precisión el estado métrico y radiométrico del edificio.

Para el desarrollo de este proyecto hemos probado con tres técnicas diferentes: láser escáner, fotogrametría y videogrametría. A pesar de la naturaleza distinta de las mismas, constituyen una posibilidad de documentación gráfica y llevan consigo una serie de características técnicas asociadas que se expondrán a continuación.

2.1 El Láser escáner

El láser scanner es un sensor activo de naturaleza topográfica que barre un área definida previamente por el software asociado, enviando ondas electromagnéticas al objeto. Estos instrumentos topográficos están basados en el mismo principio de los medidores electrónicos de distancia (EDM; Electronic Distance Measurements). La longitud de onda, entre otras variables, definirá su comportamiento, el alcance y lugar en el espectro electromagnético. Como ejemplo, los láser escáneres utilizados en los trabajos arquitectónicos tienen un ámbito de longitud de onda entre los 492nm y los 577nm aproximadamente.

Los sistemas de medición utilizados en arquitectura (LRF; Laser Range Finder) son de medio alcance y funcionan utilizando dos tipos de técnicas: Ondas de Pulso (PW; Pulsed Waves) y Ondas Continuas (CW; Continuous Wave). El caso de las Ondas de Pulso (PW-LRF) son determinadas midiendo directamente el tiempo de vuelo (TOF).



Figura 2. Nube de puntos del láser escáner Faro Focus 3d con color de la reflectancia.

Para las ondas continuas se realiza la modulación de las mismas, determinando las distancias a partir del cálculo de la diferencia de longitud de onda entre la emitida y la recibida, ya que la fase entre la emitida y la señal recibida modulada es proporcional a la diferencia de fase entre ambas. Si se aumenta la modulación de frecuencias se obtiene más resolución, en cambio si la modulación es más baja se utilizaría para obtener mejores resultados en mayores profundidades de campo. Una característica común en los láser escáneres (LRF) es su habilidad para obtener valores de mediciones de distancias y datos de la reflectancia (intensidad de la señal recibida).

En nuestro caso hemos utilizado el Faro Focus 3d, que es un escáner de naturaleza CW con un alcance máximo de 120m en condiciones favorables, y una precisión de 2,2 mm en una distancia de 25 metros. El escáner barre automáticamente el área definida por el usuario en un dispositivo táctil integrado en el sistema. Para el escaneo de la zona de pruebas se realizaron simplemente dos barridos desde posiciones semifrontales con el fin de capturar la mayor superficie posible de las columnas del templo.

2.2 Fotogrametría automática

El software fotogramétrico utilizado en este proyecto ha sido 4e (www.4-e.es). Este software comercial permite realizar una nube de puntos fotorrealista con un alto nivel de automatización pero, a diferencia de los softwares de modelado 3d vía web, 4e tiene una precisión muy alta dada la utilización de ecuaciones fotogramétricas.

El primer proceso que realiza 4e es la autocalibración. Ésta permite utilizar las mismas imágenes del objeto escogido, así que no necesita imágenes especiales (patrones de calibración o similares) para calibrar: la calibración se realiza automáticamente utilizando los conocidos parámetros de Brown (BROWN, 1971).

En el proceso de identificación de puntos homólogos, 4e utiliza una aproximación piramidal a diferentes resoluciones y escalas con el fin de agilizar la identificación de puntos homólogos y asegurar que son los correctos, al verse reducido cada vez más el área de posible encuentro (HIRSCHMÜLLER, 2008).

Para cada conjunto de imágenes se escoge una imagen principal y un conjunto de asociadas. Esta imagen principal funciona acotando el área de generación de nube de puntos. Por tanto,

para crear un modelo 3d completo de un objeto, se necesitarán múltiples imágenes principales y asociadas. Este proceso, en este ejemplo concreto, se automatizó eligiendo una imagen principal y las dos anteriores y posteriores como asociadas. Esta secuencia se siguió hasta el final de las imágenes.

Producto de esta correlación de imágenes se genera un mapa de profundidades por cada imagen principal elegida. La generación del mapa de profundidad es resultado de la proyección de cada punto 3d sobre el plano establecido en la imagen principal.

En términos prácticos, se han utilizado 14 imágenes para digitalizar el frontal del Templo de Diana, empleando una cámara Reflex Nikon D90 de una resolución de 4288x2848 pixeles. Las imágenes se han introducido en el software 4e y, dado que todos los procesos han sido automatizados, no ha habido intervención humana en el proceso. El resultado ha sido una nube de puntos (figura 3) que analizaremos posteriormente.



Figura 3. Nube de puntos con color (sin editar) procedente del software fotogramétrico 4e.

2.3 Videogrametría

De forma muy general entendemos la videogrametría como la técnica que extrae información métrica a partir de un video. Concretamente, en nuestro caso práctico, se obtiene una serie de puntos en 3d con color de forma totalmente automática (figura 4).

La metodología utilizada para este método videogramétrico está basada en formulación y algoritmos propios de la visión computacional (SNAVELY et alii, 2006). El video es una sucesión de frames (normalmente de 25 a 30 frames por segundo) y el frame tiene una configuración muy similar a una imagen fotográfica, por lo que el tratamiento puede ser similar en algunos casos. La pérdida de algunos metadatos propios de la imagen fotográfica es uno de los principales elementos diferenciadores.

La metodología empleada para la consecución de puntos viene definido primeramente por una primera identificación y orientación de los puntos a través del operador SHIFT (LOWE, 2004). La utilización de estos puntos es un primer paso para el

cálculo de la matriz fundamental F que es la matriz que nos relaciona las imágenes entre sí (HARTLEY y ZISSERMAN, 2004).

$$x^T F x = 0$$

Para un conjunto de n puntos encontrados, podemos obtener las siguientes ecuaciones lineales:

$$\begin{bmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_n x_n & x'_n y_n & x'_n & y'_n x_n & y'_n y_n & y'_n & x_n & y_n & 1 \end{bmatrix} f = 0$$

El rango de la matriz determinante A debe ser al menos de 8 puntos y si éste es exactamente 8, la solución será única, pudiéndose solucionar por métodos lineales.

A continuación se le aplica el algoritmo RANSAC a los puntos hallados. Este algoritmo responde al siguiente problema: dado un conjunto de puntos 2D, debemos encontrar el método matemático para hallar la línea que minimiza la suma del cuadrado de las distancias ortogonales (regresión ortogonal), siempre sujeto a la condición de que no serán válidos los puntos desviados desde esta línea más de una determinada unidad t .

Y utilizando los puntos resultantes del algoritmo RANSAC, se le aplica el algoritmo de los 8 puntos que diseñado originalmente por Longuet-Higgins en 1981 (LONGUET-HIGGINS *et alii*, 1994), supone el método más simple de calcular la matriz fundamental a través de un conjunto lineales de ecuaciones ajustadas por mínimos cuadrados.

Observamos una serie de puntos en el objeto en 3D ($k=1\dots m$) que son visualizados desde una serie de n cámaras cada una asociada a una matriz proyectiva ($i=1\dots n$). Las coordenadas homogéneas (x_i) de los puntos proyectados sobre las imágenes se relacionan con los puntos objeto a través de matrices y da lugar a la reconstrucción sin tener, a priori, ningún conocimiento sobre la escena ni sobre las cámaras.

Dada la información matricial de las cámaras y sus correspondientes puntos x_i , el cálculo de los puntos intersecados se denomina *triangulación*. Este proceso de triangulación se realiza compensando los parámetros internos de cada cámara con el ajuste de haces realizado mediante las ecuaciones de Levenberg-Marquardt (NOCEDAL y WRIGHT, 1999). Una vez compensados se utiliza la Transformación Lineal Directa (DLT) para el cálculo de las coordenadas terreno, prescindiendo de la escala y de la georreferenciación.

3. Resultados

La naturaleza de los resultados de las tres técnicas enunciadas anteriormente (láser escáner, fotogrametría y videogrametría) son muy diferentes a pesar de haberse presentado todos los resultados en formato de nube de puntos. Debido a la metodología del sistema de captura de información, los resultados técnicos se diferencian por lo siguiente:

- **Láser escáner.** La resolución de la nube de puntos es alta y muy homogénea debido a la forma de emisión y recepción del láser. La precisión varía proporcionalmente según la distancia del punto respecto al sensor y el ángulo de incidencia sobre el objeto. El color RGB (no la reflectancia) del mismo es bastante

deficiente debido las limitaciones propias de la cámara interna del escáner.



Figura 4. Videogrametría. Nube de puntos con color (sin editar) procedente del los algoritmos a través de los frames del video.

- **Fotogrametría.** La resolución de 4e es muy homogénea, al contrario de otros softwares de fotogrametría que basan la resolución únicamente en el encuentro de puntos homólogos. La metodología del software 4e crea un punto para cada pixel de la imagen, obteniendo así una nube de puntos homogénea. La precisión de la aplicación es subpixel. El color del punto es, con notable diferencia, el mejor de los casos aquí comparados.

- **Videogrametría.** Con el método empleado la resolución no es homogénea y la precisión no es suficiente para una aplicación profesional con requerimientos métricos, aunque procesando los datos podríamos tener un modelo visualmente aceptable.

Las condiciones de la toma en las técnicas expuestas también varían notablemente, de forma que el láser escáner debe posicionarse en distintos *scans* o posicionamientos y esperar un determinado tiempo a que éste escanee el objeto. En el caso de la fotogrametría el trabajo de campo es ínfimo pero debemos hacer notar que los resultados automáticos carecen de escala, algo que el láser escáner sí contiene. Igualmente la videogrametría es un proceso muy rápido pero, además de carecer de escala, la precisión y la resolución es deficiente para trabajos profesionales, aunque sí podría tener una utilidad en el campo de la visualización.

4. Conclusiones

La fusión e integración de las técnicas automáticas descritas en este documento, descubre la métrica de uno de los monumentos más importantes y relevantes de la ciudad emeritense. En este sentido, la fotogrametría y, por ende, la videogrametría, están tomando posición como técnica del futuro. y, en el caso de la fotogrametría, del presente, quizá por delante del láser escáner para algunas aplicaciones. El Láser escáner requiere más tiempo en la captura, siendo su movimiento más limitado además de ser un instrumento notablemente más caro. Son destacables los resultados de la nueva aplicación fotogramétrica 4e, por su gran nivel de automatismo, calidad geométrica y radiométrica de los resultados.

	Láser escáner	Fotogrametría	Videogrametría
Automático	si	si	si
Tiempo de Campo	35 minutos	4 minutos	3 minutos
Tiempo Postproceso	0	3 horas	40 minutos
Precisión	4mm	7 mm	90-120mm

Tabla1. Los resultados son producto de la extrapolación de informes de fabricantes de los sistemas a nuestro caso, considerando distancia al objeto, redundancia y condiciones de la toma. Esta extrapolación contiene elementos empíricos. [Furukawa y Ponce, 2009; www.4-e.es; www.faro.com].

La comparación de los resultados obtenidos con las técnicas aplicadas nos permite establecer la preferencia de empleo de las técnicas mencionadas, en cuanto a la adecuación a los objetivos que se buscan y la información que aportan dependiendo de su uso posterior.

La fotogrametría y, en el futuro la videogrametría, se constatan como una técnica emergente y relevante en la investigación

arqueológica. Constituyen un soporte para la puesta en valor y difusión del Patrimonio, tanto en referencia a elementos arqueológicos excavados o aún en fase de excavación, como en aquellos que, como en el caso del Templo de Diana, siempre han formado parte de la ciudad de forma visible y han permanecido como un elemento fosilizado del urbanismo romano hasta la época contemporánea.

Esta documentación gráfica constituye una forma de conservación del Patrimonio y, dada su precisión, una herramienta para la investigación del edificio y sus cambios a lo largo del tiempo, que quedan reflejados como cicatrices en sus paramentos y estructura interna.

La automatización de los procesos, simplifica su uso y democratiza su naturaleza. Perspectivas de futuro respecto a estas herramientas, nos permitirán visualizar una biografía gráfica en 3D del edificio, concebible éste como un documento arqueológico, creado a partir de la extensa documentación gráfica conservada del mismo y que aporta una lectura diacrónica de la construcción y sus metamorfosis.

Bibliografía

- ÁLVAREZ MARTÍNEZ, J.M. y NOGALES BASARRATE, T. (eds.) (2003): *Forum Colonia Augustae Emeritae. Templo de Diana*. Mérida.
- AYERBE, R; BARRIENTOS, T; PALMA, F. (eds.) (2009): *El foro de Augusta Emerita. Génesis y evolución de sus recintos monumentales*. Anegs de Archivo Español de Arqueología LIII. Mérida.
- BROWN, D. C. (1971): "Close-range camera calibration", *Photogrammetric Engineering* 37(8), pp. 855-866.
- DEL PINO ESPINOSA, B. (en prensa): "Memoria y percepción del patrimonio a través de la imagen: el Templo de Diana de Mérida", en *Proceedings of the 1st International Conference on Best Practices in World Heritage: Archaeology, Menorca, Spain, 9-13 April 2012*.
- FURUKAWA, Y. y PONCE, J. (2009): "Accurate, Dense, and Robust Multi-View Stereo", en *IEEE Xplore - Pattern Analysis and Machine Intelligence*.
- HARTLEY, R. I., y ZISSERMAN, A. (2004): *Multiple View Geometry*. Cambridge University Press, Cambridge, UK.
- HIRSCHMÜLLER, H. (2008): "Stereo Processing by Semi-Global Matching and Mutual Information", en *IEEE TPAMI Pattern Analysis and Machine Intelligence*, Vol2, páginas 328 - 341
- KAGAN, R.L. (1986): *Ciudades del Siglo de Oro. Las vistas españolas de Antón Van den Wyngaerde*, Madrid.
- LONGUET- HIGGINS, H. C.; WEBBER, B.; CAMERON, W.; BUNDY, A.; HUDSON, R.; HUDSON, L.; ZIMAN, J.; SLOMAN, A. et al. (1994) "Artificial Intelligence and Musical Cognition [and Dicussion]". *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 349 (1689): 103.
- LOWE, D. (2004): "Distinctive image features from scale-invariant keypoints", en *Int. J. of Computer Vision* 60, 2, 91.110.
- MATAS, J, CHUM, O; URBA, M, y PAJDLA, T. (2002): "Robust wide baseline stereo from maximally stable extremal regions", en *Proc. of British Machine Vision Conference*, pp. 384-396.
- NOCEDAL, J. y WRIGHT, S. J. (1999): *Numerical Optimization*. Springer Series in Operations Research. Springer-Verlag, New York, NY.
- PIERROT-DESEILLIGNY, M. y PAPANODITIS, N. (2006): "A multiresolution and optimization-based image matching approach: An application to surface reconstruction from SPOT5-HRS stereo imagery", en *LAPRS vol XXXVI-1/W41 in ISPRS Workshop On Topographic Mapping From Space (With Special Emphasis on Small Satellites)*, Ankara, Turquie, 02-2006.
- SNAVELY, N, STEVEN, M. SEITZ, SZELISKI, R. (2006): "Photo tourism: Exploring photo collections in 3D," en *ACM Transactions on Graphics (SIGGRAPH Proceedings)*, 25(3), 835-846.

Páginas webs utilizadas: www.4-e.es / www.faro.com / <http://labs.autodesk.com>