

Modelado 3D en patrimonio cultural por técnicas de *structure from motion*

José Manuel Pereira Uzal **01**

El actual desarrollo de las técnicas de fotogrametría, gracias a la evolución en el campo de la visión por computador y en particular las estrategias de *structure from motion* (SfM), se nos plantea como una nueva generación de herramientas altamente polivalentes y accesibles al profesional de la documentación o protección del patrimonio, gracias a los bajos costes en implementación. A lo largo de este artículo se explora la evolución de las tecnologías 3D, así como los fundamentos del SfM y el potencial uso de éste en las tareas de documentación de bienes culturales en general, o alteraciones en particular.

Palabras claves

Fotogrametría | Imagen digital | Patrimonio histórico | Preservación patrimonio cultural | Structure from motion (SFM) | 3D |

3D modelling in cultural heritage by structure from motion techniques

José Manuel Pereira Uzal **01**

The evolution of photogrammetric techniques thanks to the development in the computer vision techniques and strategies as the structure from motion (SFM). This strategies are a new generation of highly versatile and accessible tools for the professional in documentation and heritage preservation, thanks to low implementation costs. Throughout this article, we will study the evolution of 3D technology as well the fundamentals of SFM and the potential use of this strategies in documentation tasks in cultural heritage how a way for the recording of global or particular damage.

Keywords

Photogrammetry | Digital image | Cultural heritage | Heritage preservation | Structure from motion (SFM) | 3D |

URL de la contribución <<http://www.iaph.es/phinvestigacion/index.php/phinvestigacion/article/view/12>>

INTRODUCCIÓN

La inquietud del ser humano por representar las tres dimensiones es casi tan antigua como las propias manifestaciones artísticas de éste, y aunque objeto de estudio desde la época clásica al Renacimiento, sería en 1840 cuando Charles Wheatstone inventa el estereoscopio, capaz de recrear con una aparente sencillez la sensación de profundidad de pares de dibujos o figuras donde leves diferencias entre ambas desencadenaban la sensación de profundidad al ser procesadas por nuestro sistema de visión. En torno a 1858, años después de la invención del daguerrotipo, dicho proceso fotográfico comienza a presentarse como un medio ideal para la recreación de escenas estereoscópicas, las cuales ganarán relevancia como herramienta para el estudio de diferentes disciplinas. Sería precisamente por estos mismos años cuando el arquitecto alemán Albrecht Meydenbauer (ALBERTZ, 2002) comienza a forjar el concepto y técnicas de la fotogrametría como herramienta para el estudio de las propiedades geométricas de objetos y escenas a partir de fotografías.

Los logros de Meydenbauer no sólo deberían quedarse en el desarrollo de la fotogrametría, sino que podríamos también considerarlo uno de los pioneros de la documentación gráfica del patrimonio con fines de su preservación, ya que entre sus objetivos se encontraba el uso de las imágenes fotográficas con el fin de poder preservar la geometría de edificios o monumentos y proceder a su reconstrucción en caso de catástrofe.

En años sucesivos los principios de la fotogrametría serían aplicados al desarrollo de todo tipo de dispositivos como el estereoautógrafo o el fotocartógrafo, inspirado precisamente en el procedimiento del anaglifo patentado en 1891 por Louis Ducos du Hauron. Las imágenes de anaglifo o estereoscopia por anaglifo son las formadas por los conocidos pares de imágenes coloreadas en rojo y azul. A través de unas gafas con gelatinas de dichos colores –y gracias a los fenómenos de síntesis aditiva de nuestra visión– se neutralizan parcialmente dichas dominantes y se percibe una imagen tridimensional en blanco y negro.

Tras décadas de evolución, hoy en día la fotogrametría, ya sea de objeto-lejano –como puede ser la fotogrametría aérea– o de objeto-cercano –la orientada al estudio de escenas u objetos próximos al observador–, ha pasado de complejos dispositivos mecánicos a cómodas herramientas informáticas. Algunas de ellas, como el conocido Photomodeler de la empresa Eos Systems, lleva en la actualidad ya más de 20 años ofreciendo soluciones al campo de la fotogrametría como estrategia para el modelado 3D y estudio geométrico de objetos y escenas cercanas.

Sin embargo en las últimas décadas, los avances en visión artificial han lanzado al mercado nuevas propuestas de modelado 3D como el *structure from motion* (SfM), basado en el fenómeno por el cual los sistemas de visión de seres humanos o animales pueden reconstruir estructuras tridimensionales a partir de imágenes 2D proyectadas en la retina gracias al movimiento de estas estructuras respecto al observador o de éste, respecto a dichas estructuras. Este fenómeno también puede ser descrito por el término de profundidad cinética, quizás más usado en los ámbitos de la percepción visual, mientras que el SfM aparece asociado con frecuencia al ámbito de la visión computacional. En esta línea, productos comerciales como Photoscan de Agisoft o 123Catch del gigante Autodesk permiten el modelado y estudio de geometrías 3D de forma muy intuitiva y en pocos pasos.

Aunque ya existían trabajos en torno al SfM en los años 80, será en esta última década, gracias en parte al aumento de la potencia de cálculo de los procesadores con varios núcleos, cuando comienzan a surgir herramientas consistentes capaces de resolver modelos 3D complejos con una cierta eficacia y agilidad. Nos referimos a aportaciones como las de Changchang Wu con Multicore Bundle Adjustment (CHANGCHANG; AGARWAL; CURLESS et ál. 2011) y autor de la herramienta libre VisualSFM. Este software también se nutre de los trabajos de Yasutaka Furukawa –ingeniero de Google Maps–, desarrollador de las herramientas CMVS/PMVS encargadas de generar nubes de puntos densas a partir de grandes colecciones de imágenes; o de los trabajos de Noah Snavely, que dieron lugar a los proyectos libres Bundler, basados en el algoritmo clásico de Levenberg-Marquardt (LOURAKIS; ARGYROS, 2005); y de los trabajos de Yasutaka. Todos estos proyectos han creado un importante sustrato con las aportaciones de un buen número de algoritmos y herramientas libres para el modelado 3D a partir de imágenes 2D disponibles para cualquier usuario de forma abierta.

Mientras que la fotogrametría exigía de una caracterización previa de las cámaras a usar para la toma de imágenes, y en muchos casos una cierta planificación en dichas tomas, el SfM se vuelve más espontáneo no siendo necesario ningún tipo de planificación, ni calibración de cámaras. Siendo posible utilizar incluso imágenes tomadas con cámaras diferentes en momentos distintos, eso sí, sigue siendo necesario buscar zonas de solape, o áreas de nuestros objetos o escenas que estén contenidas en diversas imágenes.

Por tanto, hablar de SfM o fotogrametría hoy en día desde un punto diferenciador (MUNDY, 1993) resulta complejo y confuso, ya que con frecuencia la mayor parte de herramientas de fotogrametría acabaron por incorporar procesos de SfM con el fin de automatizar ciertas rutinas y liberar al usuario de las tediosas tareas de localizar manual-

mente los puntos comunes entre las diferentes imágenes, pudiendo utilizar grandes cantidades de imágenes para describir con el mayor detalle posible la geometría de objetos u escenas.

Las posibilidades de las nuevas herramientas basadas en el SfM, o en modernas técnicas de fotogrametría, nos aportan un excelente medio para la documentación gráfica 3D (STANCO, 2011; WULFF, 2010). Por su accesibilidad desde el punto de vista económico así como facilidad de implementación o manejo constituyen una alternativa a métodos de modelado 3D basados en tecnología láser (*laser imaging detection and ranging*, LIDAR) o luz estructurada. Estos últimos no sólo se presentan como inaccesibles a muchos profesionales por lo costes de su adquisición, sino también por lo complejo de su manejo y adaptación a los flujos de trabajo propios de la documentación o conservación de bienes patrimoniales, y en muchas ocasiones son poco comprendidos desde el punto de vista de la topografía.

METODOLOGÍA

El resultado de un proceso de fotogrametría o SfM es en primera instancia una nube de puntos discreta, la cual puede variar en densidad en función de las correspondencias detectadas entre imágenes. En los procesos tradicionales de ajuste manual de puntos comunes entre imágenes, estas nubes no estaban especialmente pobladas; sin embargo con SfM dichas nubes pueden llegar a tener miles de puntos. Estas nubes discretas, o poco densas, son el resultado de lo que se conoce como ajuste Bundler. Un concepto íntimamente ligado a la fotogrametría y que permite realizar la recolocación espacial tanto de una serie de puntos coincidentes entre imágenes como de las posiciones de las cámaras que tomaron dichas imágenes respecto a la escena. Dichos puntos pueden ser establecidos manualmente, o en el caso del SfM son detectados automáticamente mediante lo que se conoce como el proceso SIFT o *scale-invariant feature transform* (LOWE, 1999; MAKADIA, 2007), mediante el cual se detectan puntos o características comunes entre pares de imágenes, lo que nos permite cotejar cientos de imágenes con el fin de extraer grandes cantidades de puntos comunes o puntos clave.

Aunque estas nubes discretas puedan arrojar información sobre una geometría son insuficientes para evaluar en detalle un objeto o escena, por eso el paso final de estos procesos de SfM suele ser una nube de puntos densa con millones de puntos que describen con más detalle las superficies y geometría de objetos.

Cuando hablamos de nubes de puntos densas, ya sean obtenidas por SfM o fotogrametría, nos referimos a un conjunto de vértices descritos

en un sistema de coordenadas tridimensionales tipo XYZ. Donde además de la información espacial, para el caso de SfM y algunos LIDAR, se acompaña cada vértice o punto de una descripción colorimétrica en el modelo RGB.

Esta combinación de información geométrica o espacial con datos colorimétricos resulta especialmente interesante a la hora de recopilar información descriptiva de una obra o escena. Durante los procesos de SfM dicha información colorimétrica es extraída de los píxeles de las imágenes empleadas en el proceso, por lo que si disponemos de una fiabilidad colorimétrica en dichas imágenes, ésta será trasladada a nuestra nube de puntos densa. Este fenómeno periférico al modelado 3D nos exige por tanto una especial atención a la gestión del color en las imágenes empleadas con el fin de asegurar la fidelidad colorimétrica en lo posible.

Una vez obtenida nuestra nube de puntos densa, ya poseemos un documento sobre el que podemos realizar tanto estimaciones geométricas como colorimétricas y donde además podremos conectar nuestro flujo de trabajo con tareas de virtualización con fines meramente divulgativos, donde las texturas fotorrealísticas e iluminaciones sintéticas priman sobre la fidelidad de los modelos.

A todo modelo 3D, ya sea obtenido por una u otra técnica, se le puede atribuir una escala a partir de una distancia conocida entre dos puntos que nos permita realizar mediciones en un sistema de unidades particular. Pero también lo podemos contextualizar en un sistema de coordenadas mediante la atribución de puntos de control (*ground control points*, GCP). Esta situación nos permite construir escenarios con diversos modelos y conectarlos con sistemas de información geográfica, crear modelos basados en capas y, en particular, realizar estimaciones entre diferentes nubes de puntos con herramientas del tipo CloudCompare, el cual nos permite computar las distancias entre puntos de dos nubes aparentemente similares con el fin de poner en evidencia y cuantificar, a través de mapas de falso color, las posibles diferencias globales o particulares entre modelos. Mediante este tipo de procedimientos podemos documentar y cuantificar procesos de alteración estructural, ya sea por cambios dimensionales o pérdidas de material.

Para evaluar el potencial de estas herramientas en patrimonio, en particular como medio para la documentación o evolución de alteraciones, se ha buscado un escenario geológico natural, con un substrato pétreo de roca caliza con potenciales pérdidas de material por descamación. Dichas alteraciones naturales serán aprovechadas para documentar una situación de pérdida de material.

CASO PRÁCTICO

Nuestro flujo de trabajo comienza con la toma de 6 fotografías con una cámara DSLR común de 10 Mp, las cuales serán usadas para el proceso de SfM.

A continuación, dentro de la herramienta VisualSfM, vamos ejecutando secuencialmente los pasos descritos en el apartado de metodología relativos a la construcción de un modelo 3D basado en técnicas de SfM:

1.º Se realiza la detección de características usando el algoritmo SIFT. En este punto se crea una descripción de las características relevantes de cada imagen de forma independiente a la escala de las mismas y a los aspectos técnicos de la propia imagen tales como luminosidad, color, contraste, etc. Dicha descripción se almacena en una base de datos con el fin de ser usada en el punto siguiente.

2.º A partir de las descripciones de características de cada imagen, se realiza un cotejo por pares de imágenes localizando los candidatos a mejores puntos coincidentes o próximos por cada par de tomas.

3.º Tras ubicar las características comunes entre pares de imágenes, se ejecuta el ajuste Bundler que dará lugar a la nube de puntos dispersa más o menos densa dependiendo del número de puntos comunes entre imágenes del paso anterior.

4.º Finalmente se genera la nube de puntos densa a partir de los datos SIFT y el ajuste Bundler dando lugar a nuestro modelo finalizado descrito por una nube de puntos densa de unos 552.963 puntos para un área aproximada de 50x50 cm.

Tras realizar este proceso con los dos modelos antes y después de provocar una pérdida de material (imágenes 1 a 4), ya podemos usarlos para calcular la distancia entre sus puntos con la herramienta CloudCompare.

Con el fin de poder computar dichas distancias entre dos modelos, éstos deben someterse a un proceso de “registro” mediante el algoritmo conocido como *iterative closest point* (ICP) de forma que ambos modelos pasen a ocupar un mismo espacio, a partir del cual poder computar las diferencias. Este aspecto resulta crítico ya que podríamos generar falsas diferencias por defectos en la alineación de ambos modelos (imágenes 5 y 6).

Gracias al cómputo de las distancias entre nubes y los mapas de falso color, podemos poner en evidencia con relativa facilidad aspectos no

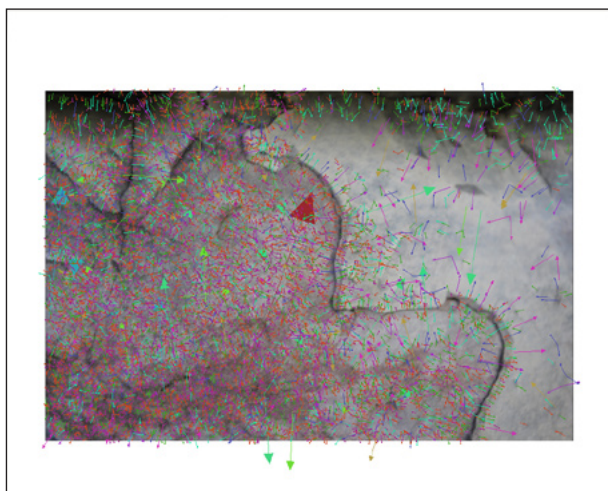


Imagen 1 |
 Descripción de las características relevantes de cada imagen

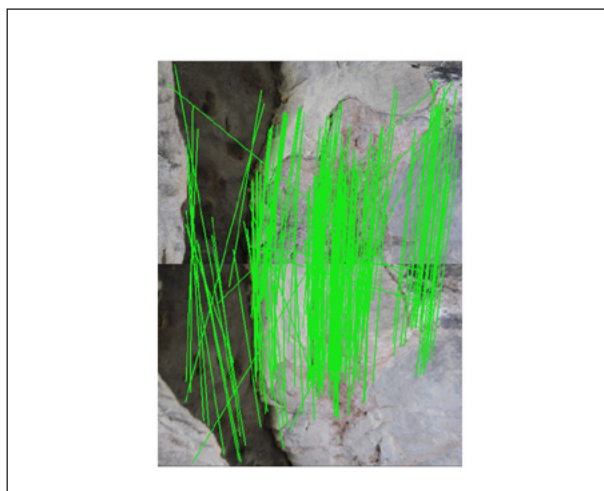


Imagen 2 |
 Una vez conocidas las características de cada imagen se pueden localizar entre pares

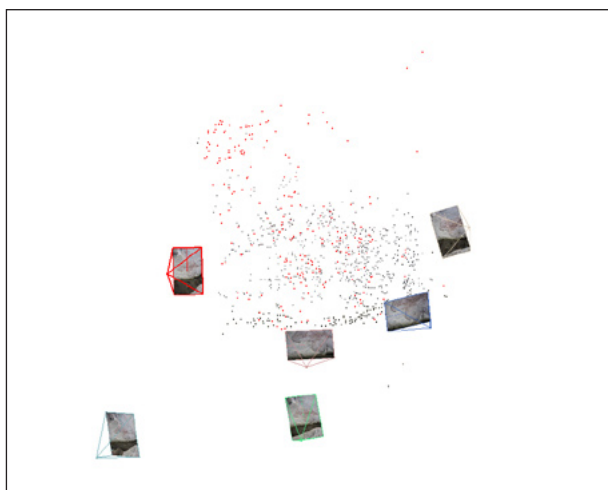


Imagen 3 |
 Aspecto de la nube de puntos dispersa tras un ajuste Bundler

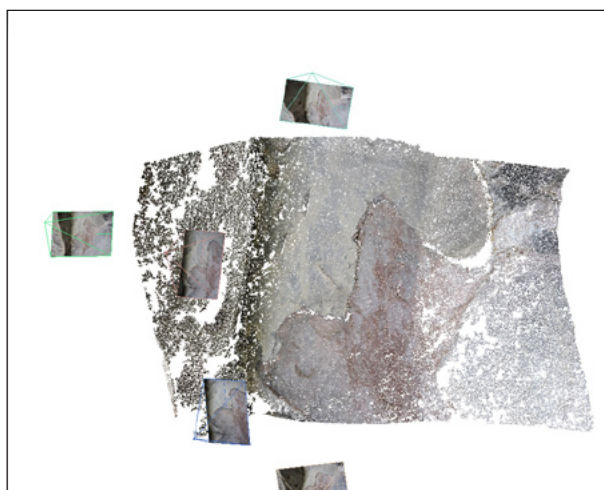


Imagen 4 |
 Nube de puntos densa con ubicaciones de las cámaras o puntos de vista de las tomas

coincidentes entre pares de modelos. Y aunque este procedimiento vaya más allá de la mera documentación o virtualización, se presenta como uno de los mayores potenciales del modelado 3D de objetos, obras o estructuras.

DISCUSIÓN

Durante el proyecto “4D arte rupestre, monitorización y conservación preventiva del arte rupestre del arco Mediterráneo de la Península Ibérica” (Ministerio de Educación, Cultura y Deporte, 13 marzo de 2013) se aplicaron de forma intensiva este tipo de metodologías, obteniendo modelos entre los 3 y 14 millones de puntos con el fin de

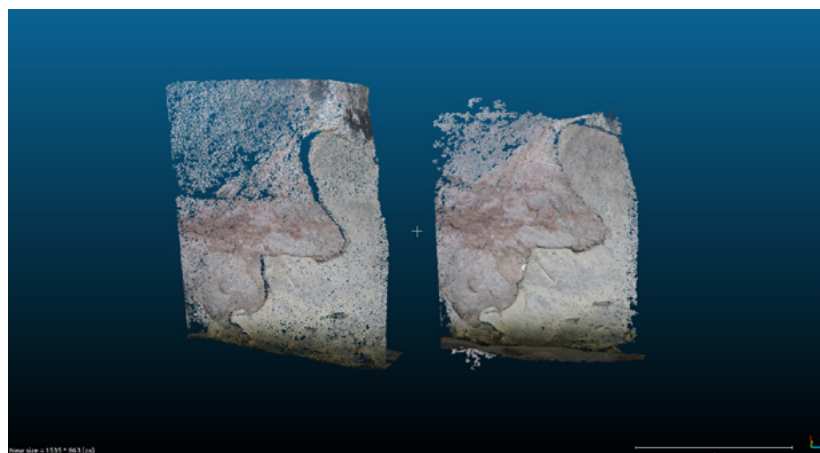


Imagen 5 |

Ambos modelos cargados en CloudCompare antes de superponerlos para su comparación.

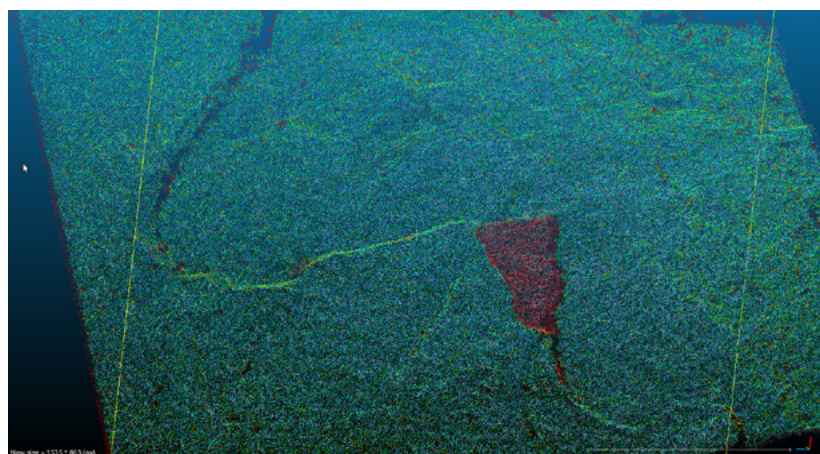


Imagen 6 |

Resultado del cálculo de distancias entre ambos modelos, donde se ha podido determinar a partir de la distancia entre puntos la región ausente en el segundo modelo.

poner en evidencia alteraciones potenciales con el paso del tiempo. Dichos modelos exigen del orden de las 60-100 fotografías para obtener modelos completos de abrigos con un importante nivel de detalle, lo cual implica una mayor demanda de recursos de cálculo y tiempo para finalizar con éxito cada proyecto.

En lo relativo al campo de la preservación digital, aunque muchos fabricantes de software y herramientas se esfuercen por aportar sus propios formatos de archivo y visores, la realidad es que actualmente disponemos de un cierto número de formatos de archivo codificados en ASCII o XML que nos pueden facilitar el acceso a dicha documentación en un tiempo futuro.

De esta forma, formatos de archivo como el *stanford triangle format* (.ply) o el *wavefront file* (.obj) nos permite describir nuestros modelos basados en nubes de puntos a través de archivos ASCII muy fácilmente comprensibles, ya que describen la información en filas y columnas bajo el formato X Y Z R G B. En una línea similar tenemos

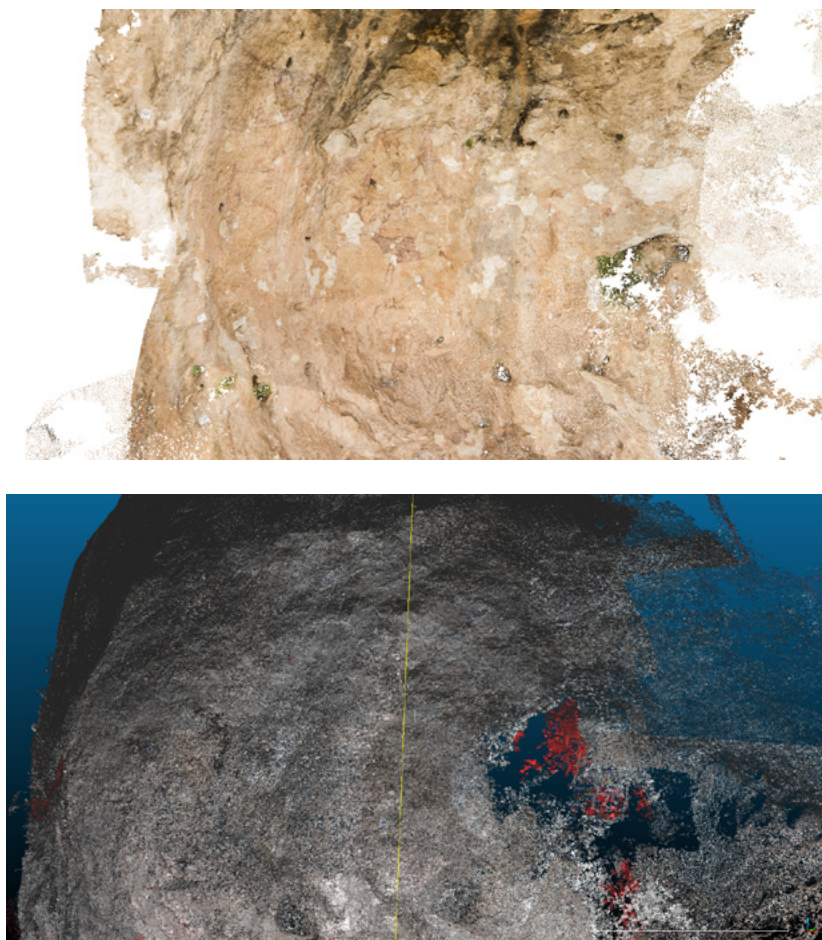


Imagen 7 |
 Modelo 3D del abrigo de la Solana de Cobachas
 6 (Nerpio, Albacete)

Imagen 8 |
 Determinación de un incremento de la
 vegetación por diferencia entre dos modelos

formatos de archivo como el X3D o el COLLADA basados en estándares abiertos y codificados en XML. Aunque las descripciones de archivo basadas en texto plano, ASCII o XML no sean las más óptimas frente a los formatos binarios, a nivel de preservación o conservación de la información digital a largo plazo son una muy buena alternativa.

Aparentemente hoy en día no resulta sencilla la reproducción de contenidos 3D en nuestros equipos por lo inusual de estos formatos de archivo, pero la realidad es bien distinta, ya que desde hace tiempo disponemos de algunos visores-editores de modelos 3D de naturaleza *open source* y multiplataforma, como el citado CloudCompare, el popular MeshLab, o el soporte ofrecido por Adobe Photoshop a modelos 3D codificados en *wavefront file*, uno de los formatos de intercambio de datos tridimensionales más populares descritos en ASCII, o *universal 3D file* una reciente estandarización proveniente de la alianza de varios fabricantes y codificado en binario, así como el soporte nativo por parte de la herramienta de vista previa de Mac Os X

a partir de su versión Snow Leopard para formatos de archivos basados en COLLADA otro reciente formato, pero en forma de estándar abierto y descrito en XML.

CONCLUSIONES

Las técnicas de SfM se proponen como un importante avance en la documentación de patrimonio, no sólo por sus importantes niveles de precisión geométrica (VERHOEVEN, 2012; ARIAS, 2006; WULFF, 2010), sino también por lo accesible (DONEUS, 2011; REUA, 2012) e inmediato de dicha técnica, cuyos niveles de detalle se ven condicionados exclusivamente por la resolución y número de imágenes a emplear.

El SfM como herramienta *open-source* nos permite, no solo acercar el modelado 3D a casi cualquier profesional de la documentación o protección del patrimonio, sino que también nos permite acometer trabajos de modelado 3D en lugares de difícil acceso para instrumentos basados en LIDAR o luz estructurada. También nos aporta inmediatez (VERHOEVEN, 2012) o espontaneidad durante las tareas de documentación, ya que exige de muy poca planificación a la hora de acometer las tomas fotográficas, frente a las tediosas planificaciones de tomas para la fotogrametría tradicional o la falta de operatividad de ciertos equipos frente a determinadas condiciones de trabajo.

Por otra parte aunque el SfM tenga como resultado nubes de puntos densas, es muy fácil de conectar con otros flujos de trabajos basados en la virtualización de objetos o escenas mediante la regeneración de superficies en mallas y texturas fotorrealísticas, así como la creación de ortofotos u ortoimágenes libres de errores de perspectiva.

BIBLIOGRAFÍA

ALBERTZ, J. (2002)

Albrecht Meydenbauer-Pioneer of Photogrammetric Documentation of the Cultural Heritage. En 18th International Symposium of CIPA. *The photogrammetric record; International Committee for Architectural Photogrammetry; Potsdam, Germany, 2001; Sep, 2002, 505.* Potsdam, Germany: Remote Sensing and Photogrammetry Society, 2002, p. 19 y ss.

ARIAS, P.; ARMESTO, J.; DI-CAPUA, D. et ál. (2006)

Digital photogrammetry, GPR and computational analysis of structural damages in a mediaeval bridge. *Engineering Failure Analysis*, vol. 14, n.º 8, December 2007, pp. 1444-1457

BRUTTO, L. M.; MELI, P. (2012)

Computer Vision Tools for 3D Modelling in Archaeology. *International Journal of Heritage in the Digital Era*, vol. 1, suplemento 1, pp. 1-6

CHANGCHANG, W.; AGARWAL, S.; CURLESS, B. et ál. (2011)

Multicore bundle adjustment. En *Procs. IEEE Conf on computer vision and pattern recognition*, 2011, pp. 3057-3064

DONEUS, M.; VERHOEVEN, G.; FERA, M. et ál. (2011)

From deposit to point cloud: a study of low-cost computer vision approaches for the straightforward documentation of archaeological excavations. *Geoinformatics*, n.º 6, 2011, pp. 81-88

HARTLEY, R. I.; MUNDY, J. I. (1993)

Relationship between photogrammetry and computer vision. *Proc. SPIE 1944, Integrating Photogrammetric Techniques With Scene Analysis and Machine Vision*, 92, September 24, 1993

LOWE, D. (1999)

Object recognition from local scale-invariant features. En *The proceedings of the seventh IEEE international conference on computer vision*. sl: IEEE Computer Society, 1999, 2 vol., pp. 1150-1157

LOURAKIS, M. I. A.; ARGYROS, A. A. (2005)

Is Levenberg-Marquardt the Most Efficient Optimization Algorithm for Implementing Bundle Adjustment? En *Proceedings of the Tenth IEEE International Conference on Computer Vision*, 2005, vol. 2, pp. 1526-1531

MAKADIA, A.; GEYER, C.; DANIILIDIS, K. (2007)

Correspondenceless Structure from Motion. *International Journal of Computer Vision*, 75 (3), 2007, pp. 311-327

REU, J. DE; PLETS, G.; VERHOEVEN, G. et ál. (2013)

Towards a three-dimensional cost-effective registration of the archaeological heritage. *Journal of Archaeological Science*, vol. 40, n.º 2, February 2013, pp. 1108-1121

STANCO, F.; BATTIATO, S.; GALLO, G. (ed.) (2011)

Digital imaging for cultural heritage preservation: analysis, restoration, and reconstruction of ancient artworks. Boca Raton, Fla.; London: CRC, 2011 (Digital Imaging and Computer Vision)

VERHOEVEN, G. (2011)

Taking Computer Vision Aloft-Archaeological Three-dimensional Reconstructions from Aerial Photographs with PhotoScan. *Archaeological Prospection*, n.º 18(1), pp. 67-73

VERHOEVEN, G.; DONEUSB, M.; BRIESEC, CH. et ál. (2012)

Mapping by matching: a computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs. *Journal of Archaeological Science*, n.º 39, 2012

WULFF, R. (2010)

Image-based 3D Documentation in Archaeology. En *DAGM 2010 32nd Annual Symposium of the German Association for Pattern Recognition* [en línea] <<http://www.dagm2010.org/pdfs/wulffdagm2010.pdf>>

Enlaces web de interés

- > **123Catch:** <http://www.123dapp.com>
- > **Agisoft:** <http://www.agisoft.ru>
- > **Bundler:** <http://phototour.cs.washington.edu/bundler/>
- > **CloudCompare:** <http://www.danielgm.net/cc/>
- > **CMVS/PMVS:** <http://www.di.ens.fr/pmvs/>
- > **MeshLab:** <http://www.meshlab.net>
- > **Photomodeler:** <http://www.photomodeler.com>
- > **Proyecto 4D arte rupestre, monitorización y...:** <http://www.4darterupestre.com/>
- > **VisualSFM:** <http://homes.cs.washington.edu/~ccwu/vsfm/>