EUROGRAPHICS Workshop on Graphics and Cultural Heritage (2026)
R. Pintus, C. Bosch, C. Schroer, P. Charalambous, I. Ciortan, D. Kotzinos, C. Andujar and D. Dininno (Editors)

M^{2}LIC: An Automated Target-Free Mobile Pipeline for Creating Relightable Models

R. Pintus[1], A. Zorcolo[1], A. Jaspe-Villanueva[2], N. Pagani[3], P. F. Serreli[3], G. Pietra[4], and E. Gobbetti[1]

[1]CRS4 - Center for Advanced Studies, Research and Development in Sardinia, Italy
[2]KAUST - King Abdullah University of Science and Technology Visual Computing Center, Saudi Arabia
[3]Independent Researcher
[4]Soprintendenza Archeologia, Belle Arti e Paesaggio per la Città Metropolitana di Cagliari e le Province di Oristano e Sud Sardegna

Abstract

Multi-Light Image Collections (MLIC) are fundamental for virtual re-lighting and surface analysis in cultural heritage. However, most standard free-form acquisition remains hindered by the need for calibration targets, specialized training, and manual processing. We present M^{2}LIC (Mobile Multi-Light Image Collection), a fully automated, target-free pipeline utilizing just two commodity smartphones, i.e., one fixed for capture, one moving for illumination. Our framework synchronizes video streams, performs frame selection to identify a small set of sharp and well-distributed images, infers light directions directly from the scene structure, and automatically generates web-ready relightable models. Validated on diverse artifacts, this rapid, open-source protocol significantly lowers the barrier to entry, enabling highly scalable MLIC digitization workflows.

CCS Concepts
• ***Computing methodologies*** → ***Computational photography; Reflectance modeling;*** *Image processing;* ***Appearance and texture representations;***

Keywords: Multi-Light Image Collections, Relightable Images, Reflectance Transformation Imaging, Polynomial Texture Mapping, Cultural Heritage, Mobile Acquisition, Target-Free Calibration, Automated Pipeline, Structure-from-Motion

1. Introduction

Multi-Light Image Collections (MLICs) capture objects from a fixed viewpoint under varying illumination, providing virtual relighting for non-invasive surface analysis, semantic annotations, and interactive web exploration in Cultural Heritage [JAP*21, PBM*25]. While automated capture systems ensure high reproducibility, their high cost, bulkiness, and size constraints restrict them to specialized labs [HHYM21]. Manual free-form capture offers a versatile, cost-effective alternative for rapid in situ documentation [LKBO24], but standard pipelines still suffer from requiring precise camera-light synchronization, obtrusive calibration spheres, and manual processing [CHI26]. Beyond high-precision models for scientific research, accessible rapid-acquisition methods are increasingly needed for remote data sharing, public presentation, and web exploration [KGM*23]. Streamlining capture by balancing metric accuracy with cost, speed, and usability lowers technical and financial barriers, enabling non-specialist practitioners to document otherwise unrecorded artifacts using commodity hardware [PB22]. Although these qualitative models may yield quantitative measurements that are not directly comparable to laboratory or calibrated setups, these qualitative models remain highly functional for many applications [PCC*10, PDC*19]: simple relighting conveys appearance and shape far better than static photography, supports interactive annotation, and reliably serves public dissemination and preliminary screening where perceptual feedback outweighs raw scientific metrics [PBM*25].

We introduce M^{2}LIC (Mobile Multi-Light Image Collection), an end-to-end pipeline democratizing MLIC acquisition. Our target-free setup uses a fixed camera and a moving camera with an attached near-axis light, easily realized with commodity smartphones. Operators record simultaneous short videos, and an automated, open-source pipeline then synchronizes the streams, extracts frames, estimates light directions, and generates web-ready relightable models without manual calibration. Primary contributions include this accessible system design, the lightweight synchronization and frame-extraction algorithm, and experimental validation on diverse cultural heritage artifacts. This paper is the short conference version of a paper accepted for publication in the ACM JOCCH Special Issue on Visual Heritage [PZJV*ar]; we refer the reader to the long version for a more detailed view of the method, additional technical details, and an extensive evaluation.

© 2026 The Authors.
Proceedings published by Eurographics - The European Association for Computer Graphics.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.2312/gch.20261061

delivered by EUROGRAPHICS DIGITAL LIBRARY www.eg.org diglib.eg.org

2. Related Work

MLIC acquisition and processing spans a wide range of computational tasks [PDC*19]. Traditional MLIC acquisition has relied on laboratory-grade instrumentation [MGW01]. To overcome these limitations, the CH community widely adopted free-form, where an operator manually positions a handheld light, inferring direction and intensity via in-scene calibration targets [PJZ*21]. Although versatile, this approach is logistically demanding, and it requires manual post-processing. The mandatory calibration targets can obstruct the artifact and are often impractical in confined *in situ* settings. Mobile solutions aim to address these limitations but introduce new constraints. Single-device approaches [YLP*17] require custom hardware attachments and computational registration, which risks compromising pixel-wise precision. Flash-on-camera methods [KHM*24] capture only backscattering signals, and necessitate additional priors [Tet24] that limit generalization. Dual-smartphone frameworks offer a promising alternative by separating the camera and light, but they remain dependent on in-scene fiducial markers [PB22]. Finally, deep-learning markerless techniques [YDPG21] are fundamentally bounded by how well the inference conditions match the training data distributions. Building on these foundations, we introduce a simple, fully automated, target-free pipeline for fast relightable model creation.

3. Method

The proposed methodology is designed to seamlessly transform raw video data into an interactive, web-ready, relightable model. The process utilizes a fixed recording device (*MLIC-Camera*), alongside a freely moving device equipped with a near-axis light attached to it (*PG-Camera*). During the acquisition phase, both cameras simultaneously record video (with audio) of the target artifact, providing the two parallel video streams that serve as the foundational input for our algorithmic pipeline. The capture routine follows a standard MLIC acquisition protocol [CHI26], requiring the operator to manually sweep the *PG-Camera* across a virtual hemisphere above the object. Since the cameras are initiated manually, it is necessary to synchronize them to ensure that any frame sampled at a specific timestamp from the *MLIC-Camera* corresponds to the same physical moment in the *PG-Camera* stream, as required for the spatial mapping of dynamic light directions. To achieve this without relying on visual frame matching or physical clapperboards, our system leverages the audio tracks embedded in the smartphone video files. The system calculates the temporal delay between the two streams by employing a cross-correlation algorithm. For the sake of image quality and computational cost, we now extract an optimized subset of corresponding frame pairs from the continuous footage. We introduce an automated, redundancy-aware frame selection strategy that filters synchronized frame pairs to identify sharp images for photogrammetric alignment. By explicitly filtering for minimal optical flow, the proposed motion-based analysis approach aims to retain frames that adhere to the point-light assumption typically required in standard MLIC processing pipelines. With the key frames isolated, a photogrammetry (SfM) pipeline spatially registers the scene, determining the static position of the *MLIC-Camera* and mapping the complete trajectory (i.e., poses) of the moving *PG-Camera* for all extracted frames, unifying both cameras within a single 3D reference frame. Because the *PG-Camera* physically carries the light source, its poses determine the 3D coordinates of the illumination source. By combining the fixed position of the *MLIC-Camera*, these tracked light positions, and the point cloud generated by the SfM pipeline, the system accurately calculates the specific incident light directions corresponding to each frame captured by the *MLIC-Camera*. Armed with the aligned *MLIC-Camera* frames and their automatically derived light directions, the system computes the final relightable image [P*25]. Ultimately, this processed representation is formatted and packaged for direct web-based visualization and interactive relighting [PBM*25]. By relying on widely available mobile devices, requiring little or no operator training, and automatically producing ready-to-view relightable models from brief capture sessions, the method aims to contribute to the democratization of MLIC acquisition and exploration, especially in the CH domain.

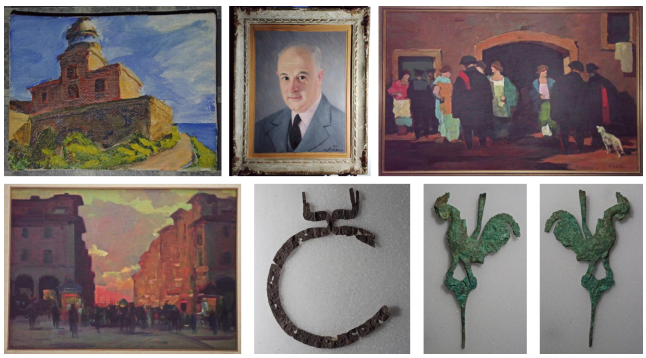


Figure 1: ***Diverse Cultural Heritage artifacts evaluated.*** *The first four are texturally complex planar artworks; the rest are 3D metallic objects from a Punic necropolis.*

4. Results

To ensure cross-platform compatibility, our Python-based pipeline utilizes free and open-source libraries and tools, such as Open3D, Pillow, NumPy, SciPy, librosa, and soundfile for data and audio processing, Meshroom [GGC*21] for 3D reconstruction, and RelightLab [P*25] and OpenLIME [PBM*25] for relightable model computation and web visualization. For capture, we rely entirely on commodity smartphones running free camera apps: a stationary Xiaomi Redmi Note 8 Pro (*MLIC-Camera*) and a handheld Xiaomi Redmi Note 10 Pro (*PG-Camera*) for continuous flash illumination.

We evaluate our method with a diverse set of CH artifacts that encompass different scales, topologies, and material properties, proving that the methodology is not only rapid and accessible but also robust across different digitization scenarios. Fig. 1 displays (from top to bottom, left to right) the small *Lighthouse* (24 x 18 cm) test painting, the *Gegeto* (40 x 31 cm), *Policemen* (98 x 61 cm), and *Sunset* (67 x 47 cm) portraits, followed by two archaeological metallic artifacts excavated from a Punic necropolis, i.e., a silver *Lamina* and a finely detailed bronze *Rooster* (both front and rear sides). The raw video recording times across all evaluated datasets range from approximately six and a half to twenty-two minutes, demonstrating the rapid and practical nature of the manual data capture phase prior to the fully automated processing pipeline.

© 2026 The Authors.
Proceedings published by Eurographics - The European Association for Computer Graphics.

To numerically assess the efficacy of our proposed methodology, we compare our target-free, automated pipeline (*Ours*) with a standard manual calibration approach (*Baseline*). The primary rationale of this evaluation is to demonstrate that eliminating the need for physical calibration targets and manual processing does not degrade the quality of the final relightable model. For this reason, the experimental acquisition setup incorporates four glossy spheres to serve as traditional targets for baseline light calibration. For quantitative evaluation, we compare two Polynomial Texture Mapping (PTM) representations obtained by those two distinct pipelines. To quantify reconstruction error, we synthesize images from the fitted PTM models under the calculated light directions and compare them against the original MLIC photographs. We evaluate the discrepancies between the real and virtually relit frames using three standard metrics: PSNR, SSIM, and LPIPS.

We evaluated our method against a manual baseline using the Lighthouse, Rooster (front and rear), and Lamina datasets, while reporting standalone absolute metrics for Gegeto, Policemen, and Sunset, which were captured natively in the field without calibration spheres. The quantitative results demonstrate that our automated, target-free pipeline performs identically to the manual, sphere-based approach. Across all baseline-compared datasets, mean PSNR values deviate by less than 0.2 dB (e.g., 28.10 vs 28.06 for *Rooster Rear*), and mean SSIM scores are practically indistinguishable, such as achieving an identical 0.9973 for *Lamina*. Average LPIPS differences also remain marginal, varying by no more than 0.006 between the two methods. For the unconstrained field captures lacking baselines, the proposed pipeline maintains strong standalone performance. It yields mean SSIM scores ranging from 0.8317 (*Gegeto*) to 0.9386 (*Sunset*) and mean PSNRs between 19.58 and 25.18 dB. The slightly lower metrics for *Gegeto* are expected and reflect the complex cast shadows introduced by its physical frame under extreme raking light, rather than a degradation of the automated acquisition protocol itself.

To illustrate *M²LIC*'s practical utility in a real-world cultural heritage investigation, we applied the pipeline to two fragile metallic artifacts currently undergoing restoration and study: the bronze *Rooster* and silver *Lamina* utilized previously for numerical benchmarks. Both objects were excavated from Is Pirixeddus, a monumental Punic necropolis in Sant'Antioco, southwestern Sardinia, containing over seventy rock-cut tombs. Recovered from elite tombs dating to the 5th–3rd centuries BC, these unique items, interpreted as a ceremonial staff terminal and a garment appliqué, lack exact parallels in known material culture and remain critical for analyzing ancient Punic funerary ideology. The fragility of these artifacts limits traditional hands-on examination. Our pipeline allows archaeologists to capture minute details, like the *Rooster*'s plumage and the *Lamina*'s embossing, using commodity devices without specialized training. Generating virtual relightable models safeguards the objects from repeated physical handling while facilitating remote comparative study. For both the silver *Lamina* and the bronze *Rooster*, the relightable models facilitate assessments of their preservation states and manufacturing techniques. Enhanced visualizations of the *Lamina* (Fig. 2a) expose original cuts, allowing scholars to reconstruct the ancient production sequence. These marks indicate the object was crafted from a larger, pre-decorated silver sheet by cutting away excess material (Fig. 2b). Furthermore,

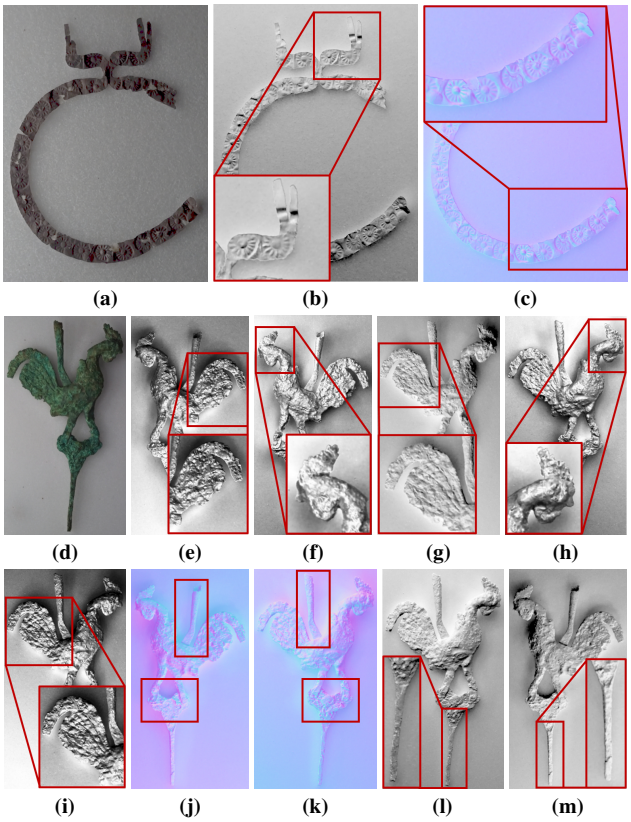


Figure 2: ***Enhanced visualizations of the silver artifact (top row) and bronze* Rooster *(center and bottom rows) revealing manufacturing and conservation details.*** *Visualizations detailing the artifact's: (a) overall conservation state; (b) cut marks indicating extraction from a larger worked sheet; and (c) repoussé decorative technique; (d) overall conservation state; (e–i) monochromatic geometry highlighting intricate plumage and crest modeling; (j, k) normal maps revealing sheet welds and assembly; and (l, m) grazing angles emphasizing the staff terminal's mounting tang.*

interactive relighting and surface normal renderings definitively confirm the use of the repoussé technique (Fig. 2c).

These visualization capabilities allow researchers to assess the artifact's preservation state (Fig. 2d) and manufacturing morphology. Isolating specific light directions and rendering monochromatic geometry without albedo reveals intricate volumetric modeling, particularly the finely rendered plumage and crest (Fig. 2e to Fig. 2i), highlighting the high level of artisanal skill involved. Furthermore, normal map visualizations explicitly reveal the weld joining the two metal body sheets and the assembly technique for the upper hollow cylindrical element (Fig. 2j and Fig. 2k). Finally, grazing angles clearly distinguish the lower portion, confirming its structural function as a tang (Fig. 2l and Fig. 2m). This evidence indicates the piece was designed for insertion into a wooden support, solidifying its interpretation as a ceremonial staff's terminal fitting.

© 2026 The Authors.
Proceedings published by Eurographics - The European Association for Computer Graphics.

5. Discussion and Future Work

Our targetless dual-smartphone workflow enables casual users to rapidly capture relightable imagery using everyday devices. Although quantitative accuracy may trail calibrated laboratory setups, the resulting qualitative models remain highly functional for interactive web viewing, public outreach, and semantic annotation [JAP*21, PBM*25], offering deeper insights into surface texture and form than static images. While demonstrated with smartphones, our pipeline scales seamlessly to professional cameras. Algorithmically, existing SfM data can be leveraged in the future to compute per-pixel light directions and compensate for radiometric falloff, enabling advanced representations like BRDFs. Absolute metric scale can be incorporated via simple reference measurements, preserving field readiness without relying on the restrictive glossy spheres required by traditional methods. Future work will balance "one-click" automation with configurable user control over formats, quality, and regions of interest. We plan to integrate rapid acquisition feedback and troubleshooting tools, empowering users to monitor the process and easily correct the pipeline if automated steps fail. Finally, we plan to conduct a comprehensive usability study involving a diverse set of CH practitioners. This will evaluate the entire ecosystem, from physical capture to software processing and visualization, to assess the accessibility and scalability of our target-free protocol and ensure it meets practical field needs.

6. Conclusions

We presented an automated pipeline democratizing MLIC acquisition in CH. Our dual-smartphone setup eliminates rigid gantries and calibration targets, lowering technical and financial barriers by leveraging everyday devices. Integrating audio synchronization and optical-flow frame selection, our pipeline uses SfM to infer 3D light positions, transforming raw mobile videos into web-ready MLICs. Quantitative evaluations confirm our automated pipeline matches traditional sphere-based baselines. Practical viability was demonstrated on real fragile archaeological artifacts. Future iterations will extend our command-line pipeline into a web interface, allowing practitioners to upload videos and receive relightable models on demand to facilitate large-scale digitization campaigns worldwide.

Acknowledgments. This research received contribution from the EXCALIBUR project funded by the European Union under the Horizon Europe Framework Program and the grant agreement N° 101233712. AJ acknowledges the contribution of KAUST under the translation research grant 3D-TURATH. EG acknowledges the contribution from Sardinian Regional Authorities under project XDATA (RAS Art9 LR 20/2015). Views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Research Executive Agency, and other funding agencies. Neither the European Union nor other granting authorities can be held responsible for them.

References

[CHI26] CHI: Cultural heritage imaging website. http://culturalheritageimaging.org, 2026. [Online; accessed-2026-03-12]. 1, 2

[GGC*21] GRIWODZ C., GASPARINI S., CALVET L., GURDJOS P., CASTAN F., MAUJEAN B., DE LILLO G., LANTHONY Y.: AliceVision Meshroom: An open-source 3D reconstruction pipeline. In *Proc. ACM Multimedia* (2021), pp. 241–247. doi:10.1145/3458305.3478443. 2

[HHYM21] HUGHES-HALLETT M., YOUNG C., MESSIER P.: A review of RTI and an investigation into the applicability of micro-RTI as a tool for the documentation and conservation of modern and contemporary paintings. *J. Am. Inst. Conserv. 60*, 1 (2021), 18–31. doi:10.1080/01971360.2019.1700724. 1

[JAP*21] JASPE VILLANUEVA A., AHSAN M., PINTUS R., GIACHETTI A., GOBBETTI E.: Web-based exploration of annotated multi-layered relightable image models. *ACM JOCCH 14*, 2 (2021), 24:1–24:31. doi:10.1145/3430846. 1, 4

[KGM*23] KHAWAJA M. A., GEORGE S., MARZANI F., HARDEBERG J. Y., MANSOURI A.: An interactive method for adaptive acquisition in reflectance transformation imaging for cultural heritage. In *Proc. ICCVW* (2023), pp. 1698–1706. doi:10.1109/iccvw60793.2023.00185. 1

[KHM*24] KAVOOSIGHAFI B., HAJISHARIF S., MIANDJI E., BARAVDISH G., CAO W., UNGER J.: Deep SVBRDF acquisition and modelling: A survey. *Computer Graphics Forum 43*, 6 (2024), e15199. doi:10.1111/cgf.15199. 2

[LKBO24] LECH P., KLEBOWSKI M., BELDYGA M., OSTROWSKI W.: Application of a stand-alone RTI measuring system with an integrated camera in cultural heritage digitisation. *J. Archaeol. Sci. Rep. 53* (2024), 104318. doi:10.1016/j.jasrep.2023.104318. 1

[MGW01] MALZBENDER T., GELB D., WOLTERS H.: Polynomial texture maps. In *Proc. ACM SIGGRAPH* (2001), pp. 519–528. doi:10.1145/383259.383320. 2

[P*25] PONCHIO F., ET AL.: Relight. http://vcg.isti.cnr.it/relight/, 2025. [Accessed 2026-03-11]. 2

[PB22] PISTELLATO M., BERGAMASCO F.: On-the-go reflectance transformation imaging with ordinary smartphones. In *Proc. ECCV* (2022), pp. 251–267. doi:10.1007/978-3-031-25056-9_17. 1, 2

[PBM*25] PONCHIO F., BETTIO F., MARTON F., PINTUS R., RIGHETTO L., GIACHETTI A., GOBBETTI E.: OpenLIME: An open and flexible web framework for creating and exploring complex multi-layered relightable image models. In *Proc. Digital Heritage* (2025), pp. 1–10. doi:10.2312/dh.20253240. 1, 2, 4

[PCC*10] PALMA G., CORSINI M., CIGNONI P., SCOPIGNO R., MUDGE M.: Dynamic shading enhancement for reflectance transformation imaging. *ACM JOCCH 3*, 2 (2010), 1–20. 1

[PDC*19] PINTUS R., DULECHA T. G., CIORTAN I., GOBBETTI E., GIACHETTI A.: State-of-the-art in multi-light image collections for surface visualization and analysis. *Computer Graphics Forum 38*, 3 (2019), 909–934. doi:10.1111/cgf.13732. 1, 2

[PJZ*21] PINTUS R., JASPE VILLANUEVA A., ZORCOLO A., HADWIGER M., GOBBETTI E.: A practical and efficient model for intensity calibration of multi-light image collections. *The Visual Computer 37*, 9 (2021), 2755–2767. doi:10.1007/s00371-021-02172-9. 2

[PZJV*ar] PINTUS R., ZORCOLO A., JASPE-VILLANUEVA A., PAGANI N., SERRELI P. F., PIETRA G., GOBBETTI E.: M2LIC: Democratizing the creation of relightable images using a targetless, free-form, automated capture workflow. *ACM JOCCH* (To appear). doi:10.1145/3848518. 1

[Tet24] TETZLAFF M.: High-fidelity specular SVBRDF acquisition from flash photographs. *IEEE TVCG 30*, 4 (2024), 1885–1896. doi:10.1109/TVCG.2023.3235277. 2

[YDPG21] YE W., DONG Y., PEERS P., GUO B.: Deep reflectance scanning: recovering spatially-varying material appearance from a flash-lit video sequence. *Computer Graphics Forum 40*, 6 (2021), 409–427. doi:10.1111/cgf.14387. 2

[YLP*17] YEH C.-K., LI F., PASTORELLI G., WALTON M., KATSAGGELOS A. K., COSSAIRT O.: Shape-from-shifting: Uncalibrated photometric stereo with a mobile device. In *Proc. e-Science* (2017), pp. 551–558. doi:10.1109/escience.2017.89. 2

© 2026 The Authors.
Proceedings published by Eurographics - The European Association for Computer Graphics.