Ancient coins' surface inspection with web-based neural RTI visualization

Leonardo Righetto[a], Enrico Gobbetti[b], Federico Ponchio[c], Arianna Traviglia[d], Michela De Bernardin[d], and Andrea Giachetti[a]

[a]University of Verona, Strada Le Grazie 15, 37134 Verona, Italy
[b]CRS4 Visual Computing, Via Ampere 2, 09134 Cagliari, Italy
[c]ISTI-CNR, Via Giuseppe Moruzzi, 1, 56127 Pisa, Italy
[d]Center for Cultural Heritage Technology, Istituto Italiano di Tecnologia, 30170 Venice, Italy

ABSTRACT

The use of neural encodings has the potential to replace the commonly used polynomial fitting in the analysis of artwork surface based on Reflectance Transformation Imaging (RTI), as it has proved to result in more compact encoding with better relight quality, but it is still not widely used due to the lack of efficient implementations available to practitioners.

In this work, we describe an optimized system to encode/decode neural relightable images providing interactive visualization in a web interface allowing multi-layer visualization and annotation. To develop it, we performed several experiments testing different decoder architectures and input processing pipelines, evaluating the quality of the results on specific benchmarks to find the optimal tradeoff between relighting quality and efficiency. A specific decoder has been then implemented for the web and integrated into an advanced visualisation tool. The system has been tested for the analysis of a group of ancient Roman bronze coins that present scarce readability and varying levels of preservation and that have been acquired with a multispectral light dome. Their level of corrosion and degradation, which in some cases hinders the recognition of the images, numerals, or text represented on them, makes the system testing particularly challenging and complex. Testing on such a real case scenario, however, enables us to determine the actual improvement that this new RTI visualization tool can offer to numismatists in their ability to identify the coins.

Keywords: Neural relighting, web-based visualization

1. INTRODUCTION

Reflectance transformation imaging (RTI)[1,2] is an extremely popular type of surface capture exploited in the Cultural Heritage domain for the interactive analysis of the objects through changes in the light source direction. The method is typically based on a direct interpolation of a set of sampled images with known illumination (Multi-Light Image Collection, MLIC), which is achieved by storing a per-pixel compact representation and a function providing the relighted color given a novel light direction. Typically, this function is fixed and controlled by few parameters, and the per-pixel representation is based on the coefficient of a polynomial fitting (Polynomial Texture Mapping, PTM[1]), or of the decomposition in Hemispherical Harmonic components (HSH[?]). Recent works showed that more realistic relighting can be obtained with a combination of Radial Basis Function interpolation and PCA decomposition[3] and that even better results are achievable by using a pixel encoding and a neural representation of the relighting provided by a neural network trained with the MLIC data.[4] This last approach, however, has never been applied for practical artwork analysis due to limitations in the efficiency of the relighting and the lack of reliable visualization software. As it is possible to see in the cultural heritage literature, while many works are exploiting RTI for surface analysis, a lot of them still use old software tools providing low-quality relighting with limited specular effects and relevant artifacts in shadowed regions.[5]

Further author information: (Send correspondence to)
A.G.: E-mail: andrea.giachetti@univr.it

The low fidelity in the material properties during the interactive relighting may be particularly bad for the analysis of surfaces with glossy parts and complex material behavior, like coins. Coins have been analyzed using RTI,[6–8] but, even in recent works, using tools allowing only polynomial or second order HSH encoding of the relightable images.

In this paper, we show that the increased relighting quality provided by neural encodings can be integrated into real-time interactive tools.

In particular, we show that it is possible to overcome the limitations of the current Neural RTI implementations by simplifying the decoder architecture keeping high-quality reflectance interpolation and optimizing the implementation of the shader allowing a real-time relight for large images on personal computers. To do this, we changed the encoder/decoder architecture to reduce the number of operations required by the decoding/relighting step, and we created a specific shader to render the relighted images that can be integrated in advanced visualization tools.

To demonstrate the effectiveness of the proposed methods, we captured with a multi-spectral light dome and visualized a set of ancient Roman bronze coins that present poor readability and different levels of corrosion and degradation, requiring careful analysis to recover useful information for their identification and study.

2. RELATED WORK

The use of a neural encoding for Reflectance Transformation Imaging was proposed by Dulecha et al.,[4] showing huge improvements in the quality of the relighted images with respect to the classic methods (PTM, HSH). The method features an asymmetric autoencoder architecture, with three fully-connected layers of fixed size (equal to $3N$ where N is the number of images) in the encoder part creating a latent representation with nine binary coefficients and a decoder with four fully connected layers of the same size transforming the latent code concatenated with the light direction in the RGB color of the pixel. A limit of the method is the complexity of the runtime decoding depending on thousands of parameters due to the fully connected architecture.

A modified NeuralRTI architecture has been proposed by,[9] replacing the encoder with a PCA-based representation in order to process a huge number of input images. However, the decoder part features one layer more than the original network and the per-pixel encoding is also heavier.

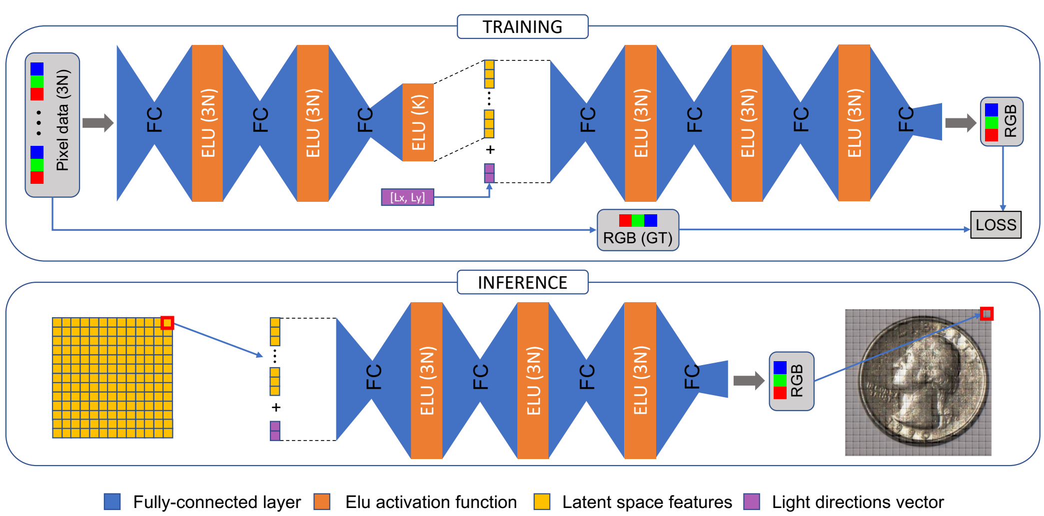


Figure 1: Original NeuralRTI scheme.[4] Top: training network with the encoder (3 layers) and the decoder (4 layers). They have the same number of parameters set to 3N. Bottom: the relighting (inference) is made with the decoder only starting from the saved latent space features (stored as image layers) and creating the RGB values for all the image pixels given an input light direction

Other methods based on neural networks have been proposed for the encoding of relightable images, even if proposed in different contexts, but not for real-time inspection in the style of RTI visualization.[10,11]

None of these methods have, however, been integrated into software tools for the interactive relighting of surfaces, that are usually employed in the Cultural Heritage domain. A list of tools of this kind is presented in an established survey paper.[5] Many of them are not public or not widely used, and most research groups in the Cultural Heritage domain still use the oldest one, RTIViewer, supporting PTM and HSH relighting and a few image enhancement filters.[12] However, some interesting, public RTI visualization tools have been recently presented, allowing a more flexible rendering and supporting different kinds of relightable image formats, e.g, MARLIE,[13] that also supports normal and BDRF representations, and OpenLIME[14]s. These tools make it possible to create rich interactive visualizations of standard and relightable images showing multiple layers and annotations.

For this reason, we developed a simplified version of Neural RTI reducing the number of required parameters. We also developed a custom shader for the interactive relighting of neural encodings, and integrated it into the OpenLIME visualization system.

3. REVISITING AND SIMPLIFYING NEURAL RTI

NeuralRTI[4] is based, in the original formulation, on the encoder-decoder structure represented in Figure 1. Both the network's sections are composed of a series of fully-connected layers, each one containing 3N parameters (where N corresponds to the number of training light directions) and provided with an ELU activation function. The encoder is required to compress input samples into feature data, the decoder takes these compressed data and a given light direction to predict a relighted image.

The network is trained end-to-end on the set of RGB values (with corresponding light direction) with the goal of minimizing the mean squared loss between predicted and measured pixel values for each specific light direction in the input data.

After the training is finished, the decoder's learned parameters and encoded feature data have to be saved. With the purpose of using this network as a web-based application, thus on a browser, particular attention should be put to data memory occupation. Feature data are basically a 3-dimensional matrix having the same resolution as input images and a number of channels equal to the number of encoding parameters, set by the

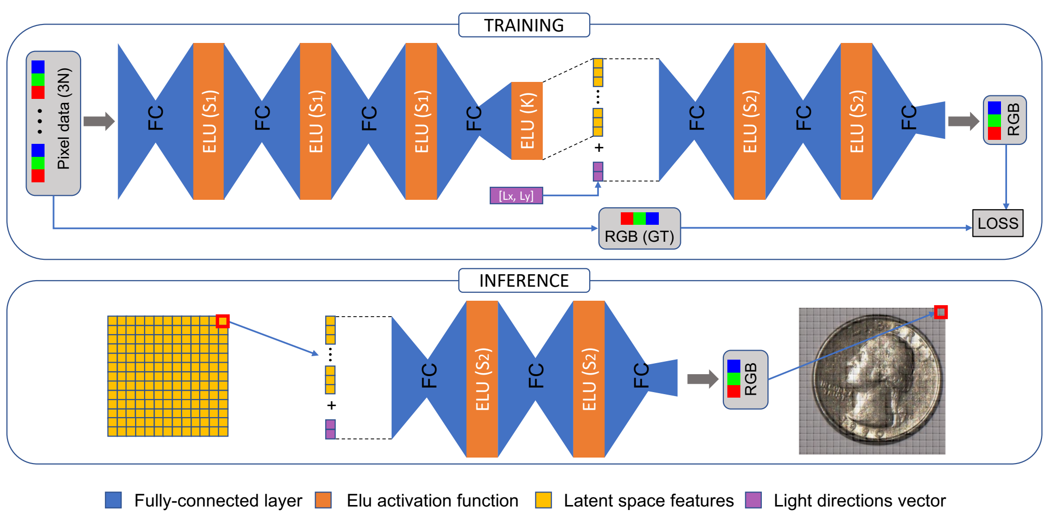


Figure 2: Modified NeuralRTI scheme. Top: training network with the encoder (4 layers) and the decoder (3 layers). They can be tuned in different ways. Bottom: inference, using the decoder only given the latent features layers and the novel light direction. Note that the size of the decoding layers can be configured (setting it as desired in the training phase and storing the information as an input for the shader).

Table 1: Comparison of average PSNR obtained on the SyhthRTI benchmark by the original Neural RTI (3L-4L) and by the classic 2nd order HSH method (values from[4]) and the modified Neural RTI with an additional encoding layer and three (4L-3L) or two (4L-2L) decoding layers.

Dataset	HSH(48)	Neural 3L-4L	Neural 4L-3L	Neural 4L-2L
SynthRTI Single	33.33	31.00	**34.60**	31.92
SynthRTI Multi	26.46	27.39	**28.80**	25.76

user. In[4] it has been shown that the method is able to give a high relighting quality with a really compact per-pixel encoding (9 parameters per pixel, that can be easily stored and distributed on the web as three RGB images, encoded as .jpg or .png files).

The main drawback of the method is that, even if the per-pixel encoding is compact and the quality high, the computational cost of the relighting is high, as it requires loading $11 * 3N + (3N)^2 + 3 * 3N$ trained weights plus $3N + 3N + 3$ trained biases and do the corresponding multiplications.

To make the decoder lighter, thus decreasing computational effort in prediction, we modified the original architecture reducing the number of parameters and related operations in the decoding phase.

In detail, we changed the number and the size of the layers of encoder and decoder, evaluating the effects of the choices on the relighting accuracy exploiting the same public benchmark (SynthRTI) used in.[4]

The modifications proposed are derived from simple observations: the first is that there are no issues in adding layers to the encoder, but it is fundamental to reduce the complexity of the decoder. For this reason, we first removed one layer from the decoder while adding one layer to the encoder checking the effects on the relighting quality. The modification in the layer structure is shown in Figure 2.

Table 1 shows that adding one more layer in the encoder and removing one layer in the decoder, keeping the layer size unchanged, the quality of the relighting measured on the benchmarks used in[4] is actually increased. Removing a second layer, instead, create a large error in the relighting of surfaces with multiple materials. The configuration with four encoder layers and three decoder layers seems therefore the optimal one.

The second observation, is that there is no need for the intermediate layers of encoder and decoder, to have a number of elements equal to the input size (3N, e.g., red, green and blue components of the N captured images of the MLIC data). In our modified model we set a constant value S_1 for the encoder layers and S_2 for the decoder, and tested the effect of varying them.

Note that our objective is to minimize the total number of the decoder parameters that increment the number of multiplications and sums to be performed in the shader. This number is obtained by summing weights (W) and biases (B) from Equation 1, where S_2 is the number of output coefficients from the FC-layers in Figure 2.

We tested the effect of changing the value of S_2 to create relightable images for the SynthRTI benchmark. We tested values proportional to the value N of the input light directions in the datasets to have values comparable with those reported in.[4] In detail, we used $S_2 \in [3N, 1N, 0.5N, 0.25N]$. The PSNR obtained in the comparison of the relighted images with the ground truth for varying S_2 on SynthRTI is reported in Table 2. Figure 3 shows that for he multi-material surfaces, that are the most relevant for practical purposes $S_2 = N =\sim 50$ works well and an increment in the layers' size does not result in increased quality.

We decided, therefore, to use this value as the default size in our implementation. However, we included the decoder layer size as a user parameter in the decoder, so that the value of S_2 can be tuned when creating the relightable images to have the desired tradeoff between relighting accuracy and decoding time. Note that, however, $S_2 = 50$ allows interactive relighting of very big images in our optimized shader.

$$\begin{aligned} W &= 11 * S_2 + S_2^2 + S_2 * 3 \\ B &= S_2 + S_2 + 3 \end{aligned} \tag{1}$$

Table 2: Average PSNR values obtained comparing original and relighted images, vs number of parameters on SynthRTI. N is the number of training lights, which for SynthRTI is equal to 49 for all objects and all materials.

Decoder config.	# of parameters	Single	Multi
3N	23964	34.60	28.80
1N	3188	34.28	28.78
0.5N	1028	33.64	28.24
0.25N	339	31.54	26.21

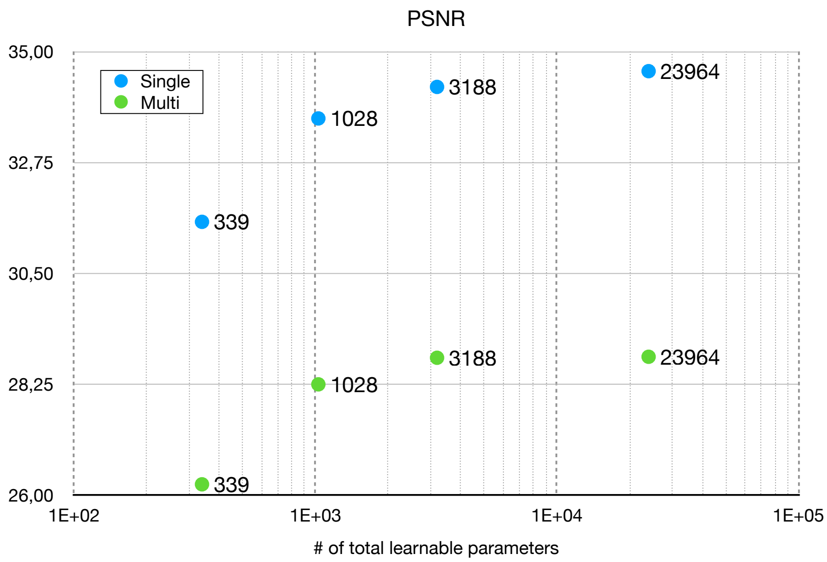


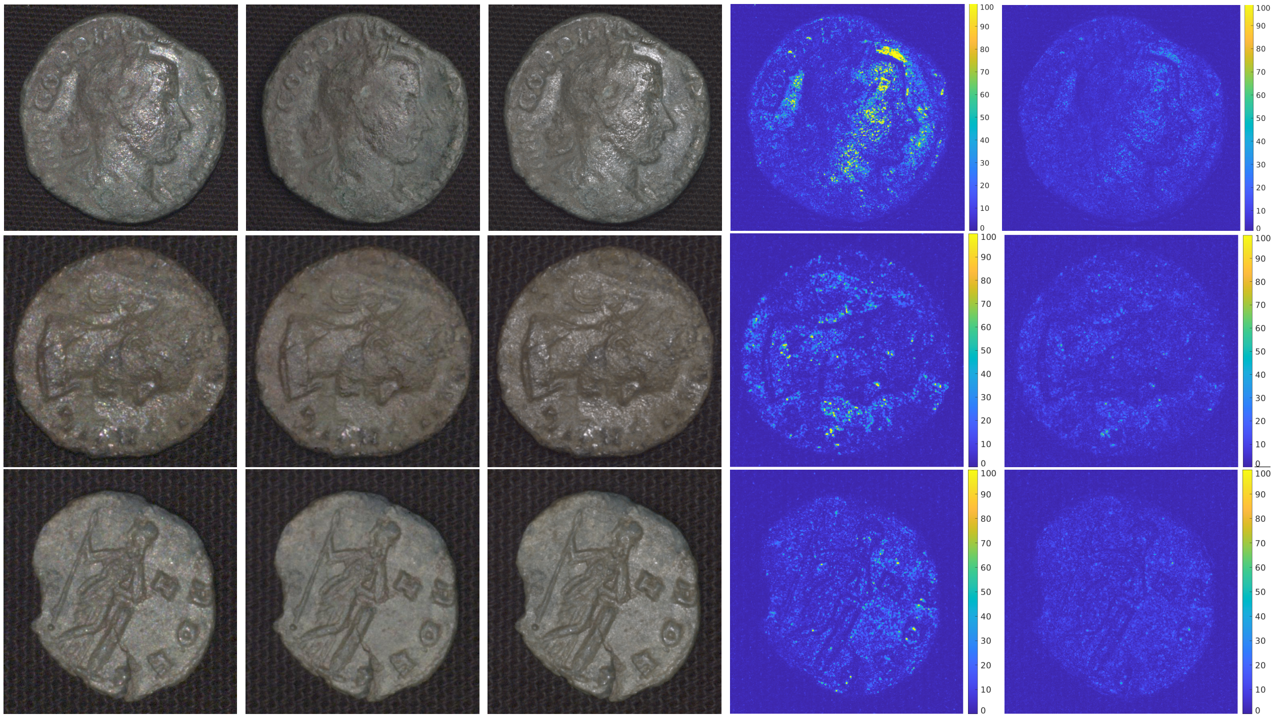
Figure 3: Average PSNR vs number of parameters on SynthRTI. The x-axis is in logarithmic scale, so to show the benefit from choosing a configuration with less parameters.

4. NEURAL RTI SHADER AND WEB VIEWER INTRGRATION

This application aims to be easy to use for cultural heritage experts, in practical use cases. The most common one is an interactive visualization of the relightable image, e.g., using the computer's mouse to dynamically change the light direction. For this reason, we focused on decreasing the complexity of neural decoding calculation, thus to allow real-time relighting.

Once the network was changed and improved, we chose a certain decoder configuration and developed a first visualization tool, in JavaScript, built upon that configuration. This way, it is possible to use the decoder on a web browser, making it easier to share it among different kinds of users. The decoding is done by exploiting a WebGL context and the OpenGL Shading Language (GLSL) to maximize the use of GPU's computing performance.

Afterward, the tool has been integrated into OpenLIME,[14] an open-source viewer useful for various kinds of visualization other than simple RTI. The application allows for interactively inspecting the captured surface with the Neural RTI quality. This makes it possible to better perceive the details of the coins thanks to the improved quality of the shadows and the specular highlights. A screenshot/video of the OpenLIME interface showing the



(a) (b) (c) (d) (e)

Figure 4: Examples of relighted images on selected coins. (a) Original image of the MLIC acquisition, captured with light coming from the top. (b) Relighting with the same light direction performed with third order HSH model (48 parameters per pixel). (c) Relighting with the same light direction performed with interactive Neural RTI (9 parameters per pixel). (d) Euclidean distance in RGB space between HSH relighting and ground truth. (e) Euclidean distance in RGB space between NeuralRTI relighting and ground truth.

Neural RTI relightable images, is shown in Figure 5. In the video it is possible to see that the user can not only inspect the surface by changing the light direction, but can also rotate the coin to make the inscriptions easier to decode, compare different layers using also a lens tool showing additional layers that can be added. In our coins' study, we added the IR and UV layers as well as the albedo and normal maps.

5. APPLICATION: RELIGHTING AND ANALYSIS OF ROMAN COINS

The novel tool for creating and analyzing relightable images has been tested on a collection of Roman coins provided by the National Museum of Aquileia (Italy) and is currently studied at the Centre for Cultural Heritage Technology (CCHT@Ca' Foscari) of the Italian Institute of Technology (IIT). The coins have been acquired with a multi-spectral light dome capturing the reflectance in the visible, IR, and UV domains.[15] Images have been pre-processed and corrected to account for the non-uniformity of the light beams using the pipeline described in.[15] Three or four coins have been acquired for each scan, placed at a distance of approximately 5cm over a dark, matte background. In this configuration the variation in the light directions within a coin region is small and the directional light assumption reasonable. We, therefore, used it, assigning to each cropped coin area the average direction derived from the calibration data within the region.

With the 36 Megapixel camera used for the acquisition (Nikon D810) the resulting resolution of the scans is then approximately 0.03mm per pixel.

Figure 4 shows a comparison of neural relighting with corresponding ground truth images and HSH (third order) relighting of three coins of the set analyzed.

The top row shows a Sestertius (Roma, 243-4 CE), representing Emperor Godianus III, and the middle row shows an Antoninianus coin of the Emperor Claudius II (Siscia, 268-70 CE), the bottom row shows another

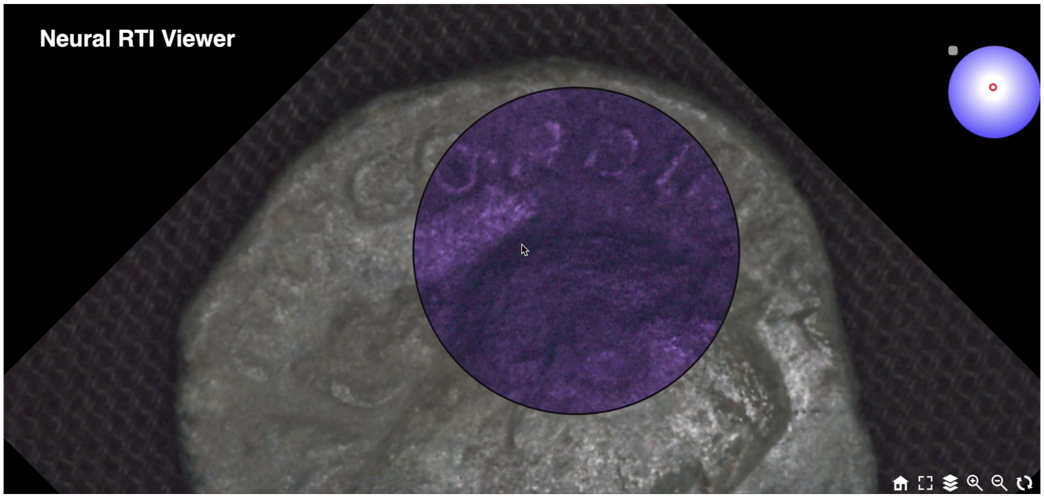


Figure 5: Video 1, showing the interactive tool for RTI-based coin analysis. The surface can be relighted from controlled directions, the view can be rotated and zoomed and the lens options can show different layers (in the examples normal map and relightable UV image). http://dx.doi.org/doi.number.goes.here

Antoninianus of Emperor Gallienus (Roma, 260-8 CE). For all the selected coins the improved relighting can provide improvements in the reading of legends and iconographic types. In the NeuralRTI relighting, it is possible to appreciate the better contrast and the highlights and shadows quite similar to the ground truth acquisition. The error maps representing the RGB distance of pixel colors reveal a large improvement in the relighting quality.

Thanks to the simplified network and the custom shader, high-quality relighting can be performed interactively within a web interface. The interface also provides real-time (> 30 fps.) performances on low-end graphic cards and supports the visualization of multiple layers, zooming, panning, and rotating. Currently the application can visualize static albedo and normal maps and relightable images captured with visible, UV and IR lights. Figure 5 shows a video where the user first selects a relighting direction, zooms in then uses the lens tool to show the normal map in selected regions. In the second part of the video, the view is rotated and zoomed to better perceive the inscription, the light direction is adjusted and finally the relightable UV image is displayed in the lens region.

The implementation of the relighting in the custom shader also enables the easy addition of other solutions to enhance interactively the visual perception of detail (gamma correction, unsharp masking, edge detection).

6. CONCLUSIONS

In this paper we have shown an improved architecture for the creation of neural relightable images that can be exploited for the interactive analysis of the surface of Cultural Heritage artifacts. We demonstrated the usefulness of the tool for the analysis of ancient coins that can feature complex material behavior and require a clear perception of the shape details. We plan as future work to quantify the advantages of the proposed method with user tests with numismatic experts and to further improve the potential impact of the method by adding specific image enhancements options simplifying the decoding of inscriptions and decorative patterns.

ACKNOWLEDGMENTS

The authors would like to acknowledge the National Museum of Aquileia (Italy) for their support and collaboration. The images published here are made available under the concession granted by the Ministry of Culture, Regional Directorate of Museums of Friuli Venezia Giulia, and are subject to specific conditions. The coins shown in this article are identified with the Italian National IDs: 579666, 587202, 582404 and are held at the Museo Archeologico Nazionale of Aquileia (UD). Any further reproduction, duplication by any means, as well as downloading and subsequent manipulation, is strictly prohibited.

This study was partially carried out within the PNRR research activities of the consortium iNEST (Interconnected North-Est Innovation Ecosystem) funded by the European Union Next-GenerationEU (Piano Nazionale di Ripresa e Resilienza (PNRR) – Missione 4 Componente 2, Investimento 1.5 – D.D. 1058 23/06/2022, ECS_00000043).

REFERENCES

[1] Malzbender, T., Gelb, D., and Wolters, H., "Polynomial texture maps," in [*Proceedings of the 28th annual conference on Computer graphics and interactive techniques*], 519–528 (2001).

[2] Earl, G., Basford, P., Bischoff, A., Bowman, A., Crowther, C., Dahl, J., Hodgson, M., Isaksen, L., Kotoula, E., Martinez, K., et al., "Reflectance transformation imaging systems for ancient documentary artefacts," *Electronic visualisation and the arts (EVA 2011)* , 147–154 (2011).

[3] Ponchio, F., Corsini, M., and Scopigno, R., "Relight: A compact and accurate rti representation for the web," *Graphical Models* **105**, 101040 (2019).

[4] Dulecha, T. G., Fanni, F. A., Ponchio, F., Pellacini, F., and Giachetti, A., "Neural reflectance transformation imaging," *The Visual Computer* **36**, 2161–2174 (2020).

[5] Pintus, R., Dulecha, T. G., Ciortan, I., Gobbetti, E., and Giachetti, A., "State-of-the-art in multi-light image collections for surface visualization and analysis," in [*Computer Graphics Forum*], **38**(3), 909–934, Wiley Online Library (2019).

[6] Palma, G., Siotto, E., Proesmans, M., Baldassari, M., Baracchini, C., Batino, S., and Scopigno, R., "Telling the story of ancient coins by means of interactive rti images visualization," *Archaeology in the digital era* , 177 (2014).

[7] Kotoula, E. and Kyranoudi, M., "Study of ancient greek and roman coins using reflectance transformation imaging," *E-conservation magazine* **25**, 74–88 (2013).

[8] Min, J., Jeong, S., Park, K., Choi, Y., Lee, D., Ahn, J., Har, D., and Ahn, S., "Reflectance transformation imaging for documenting changes through treatment of joseon dynasty coins," *Heritage Science* **9**(1), 1–12 (2021).

[9] Pistellato, M. and Bergamasco, F., "On-the-go reflectance transformation imaging with ordinary smartphones," in [*Computer Vision–ECCV 2022 Workshops: Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part I*], 251–267, Springer (2023).

[10] Li, J. and Li, H., "Neural reflectance for shape recovery with shadow handling," in [*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*], 16221–16230 (2022).

[11] Yang, W., Chen, G., Chen, C., Chen, Z., and Wong, K.-Y. K., "S ˆ3-nerf: Neural reflectance field from shading and shadow under a single viewpoint," *arXiv preprint arXiv:2210.08936* (2022).

[12] Palma, G., Corsini, M., Cignoni, P., Scopigno, R., and Mudge, M., "Dynamic shading enhancement for reflectance transformation imaging," *ACM Journ. on Computers and Cultural heritage* **3** (set 2010).

[13] Jaspe Villanueva, A., Ahsan, M., Pintus, R., Giachetti, A., and Gobbetti, E., "Web-based exploration of annotated multi-layered relightable image models," *ACM Journal on Computing and Cultural Heritage* **14**, 24:1–24:31 (may 2021).

[14] CRS4 Visual and Data-intensive Computing Group and CNR ISTI - Visual Computing Lab, "Openlime (open layered image explorer)," (2023). https://github.com/cnr-isti-vclab/openlime - accessed 2023-04-20.

[15] Ciortan, I. M., Dulecha, T., Giachetti, A., Pintus, R., Jaspe-Villanueva, A., and Gobbetti, E., "Artworks in the spotlight: characterization with a multispectral led dome," in [*IOP Conference Series: Materials Science and Engineering*], **364**(1), 012025, IOP Publishing (2018).