A Compact Representation of Relightable Images for the Web

Federico Ponchio
ISTI-CNR
Pisa, Italy
federico.ponchio@isti.cnr.it

Massimiliano Corsini
ISTI-CNR
Pisa, Italy
massimiliano.corsini@isti.cnr.it

Roberto Scopigno
ISTI-CNR
Pisa, Italy
roberto.scopigno@isti.cnr.it

ABSTRACT

Relightable images have demonstrated to be a valuable tool for the study and the analysis of coins, bas-relief, paintings, and epigraphy in the Cultural Heritage (CH) field. Reflection Transformation Imaging (RTI) are the most diffuse type of relightable images. An RTI image consists in a per-pixel function which encodes the reflection behavior, estimated from a set of digital photographs acquired from a fixed view. Even if web visualization tools for RTI images are available, high fidelity of the relighted images still requires a high amount of data to be transmitted. To overcome this limit, we propose a web-friendly compact representation for RTI images which allows very high quality of the rendered images with a relatively small amount of data required (in the order of 6-9 standard JPEG color images). The proposed approach is based on a joint interpolation-compression scheme that combines a PCA-based data reduction with a Gaussian Radial Basis Function (RBF) interpolation. We will see that the proposed approach can be adapted also to other data interpolation schemes, and it is not limited to Gaussian RBF. The proposed approach has been compared with several techniques, demonstrating its superior performance in terms of quality/size ratio. Additionally, the rendering part is simple to implement and very efficient in terms of computational cost. This allows real-time rendering also on low-end devices.

CCS CONCEPTS

• **Computing methodologies** → **Reflectance modeling**; *Graphics file formats*; *Image compression*;

KEYWORDS

Relightable images, Compact representation, Compression, Graphics for the Web, PTM, RTI

ACM Reference Format:
Federico Ponchio, Massimiliano Corsini, and Roberto Scopigno. 2018. A Compact Representation of Relightable Images for the Web. In *Web3D '18: The 23rd International Conference on Web3D Technology, June 20–22, 2018, Poznan, Poland.* ACM, New York, NY, USA, Article 4, 10 pages. https://doi.org/10.1145/3208806.3208820

Web3D '18, June 20–22, 2018, Poznan, Poland

ACM ISBN 978-1-4503-5800-2/18/06…$15.00
https://doi.org/10.1145/3208806.3208820

1 INTRODUCTION

Reflection Transformation Images (RTI) are a special type of images where the user can interactively relight the image subject, thus simulating the classical interactive inspection action of modifying the lighting to better appreciate the shape details. This type of images is increasingly used in the Cultural Heritage (CH) field because the way the light interacts with the object of interest allows disclosing important information on the surface conservation status or the constituent materials. The reflectance behavior of the materials and the perception of the fine details are very important for the study of bas-relief [Hammer et al. 2002], coins [Mudge et al. 2005; Palma et al. 2012], paintings [Giachetti et al. 2017, 2015; Padfield et al. 2005], or epigraphy [Lamé 2015].

RTI images do not encode a color for each pixel, but a function that allows computing the specific color associated to each pixel given the light direction. Therefore, the RTI encodes an approximation of the reflectance behavior of the sampled surface, and each pixel is relighted in real-time according to the light direction sets dynamically by the user. Typically, a RTI image is generated through a digital acquisition process that captures a set of images of the subject, all of them shot from the same fixed viewpoint, under varying lighting conditions. Then, these data are used to generate the per-pixel reflectance function just mentioned.

The alternative way to conduct similar visual inspections is to acquire a high-fidelity 3D model of the geometry of the object of interest and map on this 3D model the sampled and reconstructed reflectance (BRDF) of the surface. However, acquiring the geometry and estimating the reflectance properties at a very fine scale is difficult and time-consuming. This is why the use of RTI is becoming quite common in the CH domain: RTI technologies are cheaper and more effective than 3D scanning for some applications.

Recently, the RTI approaches have been extended to enable multispectral analysis of the acquired data [Giachetti et al. 2017] under different lighting conditions. In this recent work, the per-pixel function fitting is replaced by an *interpolation* approach. This allows achieving higher quality of the relighted images produced in real time, since the reproduction is much closer to the original data acquired. The main disadvantage of this interpolation-based approach is a much larger quantity of data required to generate each relighted image. This prevents to offer this method on a web-based RTI viewer, such as the ones based on standard RTI formats [Palma et al. 2012]. To overcome this limit, we propose a joint *interpolation-compression* scheme, which enables interpolation-based RTI visualization on the Web, thus requiring a reasonable amount of data to be transmitted. To achieve this goal we apply Principal Component Analysis (PCA) to reduce the amount of data required for the interpolation. This data reduction approach is similar to the PCA-based methods employed for the compression of high-resolution Bidirectional Texture Functions (BTF) [Müller et al. 2003; Sattler et al. 2003].

In the following, we will demonstrate that the basis generated for the reflection function can easily integrate the PCA compression and the Gaussian Radial Basis Function (RBF) interpolation, taking the advantages of the two approaches. This final compact representation is transmitted through the Web as a set of JPEG images. Additionally, the final interactive rendering consists in a simple weighted linear summation of coefficients that can be easily implemented in a pixel shader. The proposed approach has been validated on a set of test cases, presented in the Results Section, showing great benefits w.r.t other standard techniques.

2 RELATED WORK

Different approaches can be used to produce relightable images. Image relighting can be obtained if the surface normals and the BRDF of the object/scene to relight are known. Nevertheless, estimating a 4D BRDF is a complex task, especially in the case of complex multi-material surfaces or degraded artworks. Describing in detail the state of the art of BRDF acquisition is outside the scope of this work; interested readers can refer to [Dorsey et al. 2008]. Another way to acquire a photo-realistic reflectance field of a material can be achieved reconstructing a Bidirectional Texture Function (BTF). BTF consists in acquiring a small sample of material in different lighting and viewing conditions. In particular, for each light direction and for each viewpoint a small piece of material is acquired and a normalized texture is generated and used for successive photorealistic rendering [Schwartz et al. 2013].

An effective trade-off between full sampling of the reflectance and acquisition/rendering complexity is to limit the acquisition to a fixed viewpoint. These methods start from the selection of a specific view of interest and store the reflectance field as a per-pixel function. A pioneer work of this type is Polynomial Texture Maps (PTM) [Malzbender et al. 2001]. PTM can be seen as a particular case of BTF, where the single view of an entire scene/object is acquired under different lighting conditions.

In the last years, many variants of the PTM have been proposed to overcome the limitation of approximating the reflectance function with a simple bi-quadratic polynomial. Consequently, the more general denomination of Reflectance Transformation Imaging (RTI) emerged. More sophisticated basis functions have been studied to better represent the reflectance, like spherical (SH) and hemispherical harmonics (HSH) [Mudge et al. 2008]. Eigen hemispherical harmonic (EHSH) have been proposed as a better basis of SH in terms of controllability of the final coefficients [Lam et al. 2012]. In order to better represent certain materials, more physical representations, that separate the diffusive and the high-frequency specular components of the reflection, have been proposed [Pintus et al. 2017; Zhang and S Drew 2014]. For example, the work by Zhang and Drew et al. [Zhang and S Drew 2014] used a PTM to represent the matte part of the material and approximate the residuals of the high-frequency part using radial basis functions (RBF).

Recently, Multi-Spectral Reflectance Imaging (MS-RTI) has been proposed by Giachetti et al. [Giachetti et al. 2017]. In MS-RTI, five frequency bands (IR, UV and visible spectrum) are acquired under different lighting conditions. Those densely acquired data are stored directly in the data structure and interpolated using Gaussian RBF at rendering time. This approach allows a more accurate analysis of the reflectance data acquired over the artifact. Unfortunately, this approach is not efficient in a web-based visualization context, because we need all the input data (50-100 full resolution images) to compute the interpolation requested by the user in real-time.

A lot of work about the compression of *reflectance data* has been done in the scope of Bidirectional Texture Function (BTF) due to the large size of the data involved. The proposed solutions can be categorized in two classes [Filip and Haindl 2009]; the pixelwise BRDF methods and the linear factorization methods. The first methods aim to find a compact representation of the apparent BRDF (BRDF including self-shadowing effects) for each pixel. Some solutions for BTF with fixed view approximate the reflectance function with pixelwise Lafortune lobes [Meseth et al. 2003], generalization of the one-lobe Lafortune lobe [Filip and Haindl 2004] and its polynomial extension [Filip and Haindl 2004, 2005]. Linear factorization methods decompose the acquired data using a linear basis. Typically SH or Principal Component Analysis (PCA) are used for this purpose. Koudelka et al. [Koudelka et al. 2003] proposed to arrange each image as a vector and form a matrix factorized with SVD. Reflectance Field Factorization [Sattler et al. 2003] used no more than n PCA components to facilitate the rendering and achieved good compression ratio. Müller et al. [Müller et al. 2003] partitioned the BTF space in clusters and applied PCA to each cluster, the so-called local PCA. This allows high compression ratio and high quality rendering. Ho et al. [Ho et al. 2005] exploit blockwise PCA in the YCbCr color space.

From our knowledge, the only two works specific for RTI compression are the work by Schuster et al.[Schuster et al. 2014] and the one by Schwartz et al. [Schwartz et al. 2013]. Schuster et al.[Schuster et al. 2014] proposed a method to send RTI images to mobile devices by transmitting the luminance and the chrominance planes of each PTM compressed as JPEG (or JPEG2000) images. The idea is to use different quality for the luminance and for the chrominance. This simple approach is effective in reducing the bit rate while maintaining a good overall quality of the rendered image w.r.t the original PTM. Schwartz et al. [Schwartz et al. 2013] used Decorrelated Full Matrix Factorization (DFMF) for the efficient streaming of BTF on the Web. Many design choices of this algorithm are targeted to the BTF, for example the components are properly ordered to receive before the most important visual components.

Our approach follows the approach of some of the just mentioned PCA-based methods for BTF, [Sattler et al. 2003] and [Müller et al. 2003], in particular. In our case we do not treat a single material sample but an entire scene under fixed viewpoint. To preserve as much as possible the appearance of the acquired data, we combine the PCA with an RBF interpolation scheme, demonstrating that it is possible to integrate these two operations into a simple weighted linear sum. This guarantees, at the same time, a moderate computational burden in the pre-processing step, high quality/size ratio with respect to standard RTI techniques, such as PTM or HSH, and simple and efficient implementation of the visualization client.

3 EXISTING RTI REPRESENTATIONS

As previously stated, an RTI image is often represented as a per-pixel linear basis. We would like to introduce here some formulations that will help to better clarify our jointly compression-interpolation scheme. In the most general case we have:

$$\begin{pmatrix} R(x,y) \\ G(x,y) \\ B(x,y) \end{pmatrix} = \begin{pmatrix} f_R(x,y,l) \\ f_G(x,y,l) \\ f_B(x,y,l) \end{pmatrix} \tag{1}$$

where $(R(x,y), G(x,y), B(x,y))$ is the final color of the image pixel (x,y) and $l = (l_u, l_v)$ is the light direction vector (normalized) projected on the image plane. The function f assumes the form of a linear weighted summation:

$$f(x,y,l) = \sum_k a_k(x,y) w_k(l) \tag{2}$$

where the weights depend on the light direction, that is assumed fixed for each pixel, and the coefficients a_k describe the reflectance behavior at pixel level.

In the case of Polynomial Texture Maps (PTM) [Malzbender et al. 2001], we have two cases, LRGB PTM and RGB PTM. In the last case, we have

$$\begin{aligned} f_R(x,y,l) &= c_o^R(x,y) + c_1^R(x,y)l_u + c_2^R(x,y)l_v + \\ &+ c_3^R(x,y)l_u l_v + c_4^R(x,y)l_u^2 + c_5^R(x,y)l_v^2 \qquad (3) \\ f_G(x,y,l) &= c_o^G(x,y) + c_1^G(x,y)l_u + c_2^G(x,y)l_v + \\ &+ c_3^G(x,y)l_u l_v + c_4^G(x,y)l_u^2 + c_5^G(x,y)l_v^2 \qquad (4) \\ f_B(x,y,l) &= c_o^B(x,y) + c_1^B(x,y)l_u + c_2^B(x,y)l_v + \\ &+ c_3^B(x,y)l_u l_v + c_4^B(x,y)l_u^2 + c_5^B(x,y)l_v^2 \qquad (5) \\ R(x,y) &= f_R(l,x,y) \\ G(x,y) &= f_G(l,x,y) \qquad (6) \\ B(x,y) &= f_B(l,x,y) \end{aligned}$$

so that the light-dependent base is $w(l) = \{1, l_u, l_v, l_u l_v, l_u^2, l_v^2\}$ and the per-pixel coefficients are the coefficients of the bi-quadratic polynomial $a_k = c_k$. Since the luminance exhibits a great variation under lighting changes with respect to the chrominance, a variant called LRGB PTM with less coefficients has been proposed:

$$\begin{aligned} L(x,y,l) &= c_o(x,y) + c_1(x,y)l_u + c_2(x,y)l_v + \\ &+ c_3(x,y)l_u l_v + c_4(x,y)l_u^2 + c_5(x,y)l_v^2 \qquad (7) \\ R(x,y) &= L(l,x,y)R_B(x,y) \qquad (8) \\ G(x,y) &= L(l,x,y)G_B(x,y) \\ B(x,y) &= L(l,x,y)B_B(x,y) \end{aligned}$$

In this case the "luminance" $L(.)$ modulates a basic color (R_B, G_B, B_B) calculated during the coefficients estimation. Note that the $L(.)$ term is not a real luminance, because it includes also self-shadowing effects. Hence, a RGB PTM is characterized by 18 coefficients per pixel (6 for each color channel) while the LRGB PTM by 9 coefficients per pixel (6 for the $L(.)$ function and 3 for the base color).

If Spherical Harmonics (SH) are used to represent the lighting-dependent weights in Eq. (2) become:

$$w_k(\theta,\phi) = \begin{cases} \sqrt{2}K_l^m \cos(m\phi)P_l^m cos\theta & m > 0 \\ \sqrt{2}K_l^m \sin(m\phi)P_l^m cos\theta & m < 0 \\ K_l^0 P_l^0 cos\theta & m = 0 \end{cases} \tag{9}$$

The light direction l is represented using the azimuth ($\phi \in [0, 2\pi]$) and elevation ($\theta \in [0, \pi]$) angles. P_l^m are the associated Legendre Polynomial of degree l ($-l \leq m \leq l$) and K_l^m are the corresponding normalization factors. Hemi-Spherical Harmonics (HSH) are more used than SH, since in general the reflective phenomena happen in the upper hemisphere. They have a slightly different definitions and employ shifted Legendre polynomial [Zhang and S Drew 2014].

In other approaches, *interpolation* is used instead of linear decomposition. This is the case of the recent multi-spectral RTI representation proposed by Giachetti et al. [Giachetti et al. 2017]. The final image is reconstructed by finding the light directions closest l and interpolate them using a set of Radial Basis Function (RBF). In case of Gaussian RBF we have:

$$I(x,y) = \sum_{i=1}^{n} a_k \exp\left(\frac{\|l - l_i\|_2^2}{\sigma^2}\right) \tag{10}$$

where the $I(x,y)$ is the 5-components pixel of the multi-spectral bands acquired (IR, UV and three for the visible spectrum). In this case the light-dependent weights do not depend directly on l but depend on the closest lights direction, i.e. $w_k(g(l))$. The implementation presented in [Giachetti et al. 2017] interpolates the 5 closest light directions.

Our formulation, described in the next section, integrates interpolation (similar to the one proposed by Giachetti et al. []) with PCA-based linear decomposition to reduce the amount of data required.

4 PROPOSED REPRESENTATION

The adoption of Radial Basis Functions (RBF) for the interpolation of the N images acquired in different lighting conditions has been demonstrated [Giachetti et al. 2017; Zhang and S Drew 2014] to produce better visualization of materials and shape details. As shown in Eq. (10), the final color of a pixel for a given light direction l is given as a weighted sum of Gaussians centered on l. Indicating with $\rho = (R, G, B)$ the final color of a pixel, Eq. (10) can be rewritten as:

$$\rho(x,y,l) = \sum_{i=1}^{N} \rho_i(x,y) \exp\left(\frac{\|l - l_i\|_2^2}{\sigma^2}\right) = \sum_{i=1}^{N} \rho_i(x,y)\phi_i(l) \tag{11}$$

The parameters σ is the Gaussian dispersion. It controls the amount of smoothing between samples and needs to be adjusted to the density of the lighting sampling (for this purpose, the optimization method described in [Drew et al. 2012] can be used). In our proposal, we consider all the samples such that their contribution is not negligible. Obviously, this number depends on the value of σ.

In general, this approach is not suitable for a web-based visualization due to the size of the dataset (50-100 images) to be transmitted.

Our idea is to reduce the amount of data to be interpolated, using Principal Component Analysis (PCA), following an approach similar to the ones proposed for BTF compression.

Considering all the acquired N images together, each pixel can be expressed as a vector p with $3N$ components. By applying PCA we obtain a mean μ and a basis B of $M < N$ vectors such that:

$$p(x, y) \simeq \mu + \sum_{j=1}^{M} a_k(x, y) B_k \tag{12}$$

where the basis B is formed by the k eigenvectors with the highest eigenvalues of the correlation matrix pp^T and $a_k(x, y)$ are the coefficients to represent the vector on this basis. If we indicate with ρ_i the three color components of the i-th image, we can combine this approximation with the RBF representation (Eq. (11)):

$$\begin{aligned} \rho(x, y, l) &\simeq \sum_{i=1}^{N} \phi_i(l) \left(\mu_i + \sum_{k=1}^{M} a_{i,k}(x, y) B_{i,k} \right) \\ &\simeq \sum_{i=1}^{N} \phi_i(l) \mu_i + \sum_{k=1}^{M} a_{i,k}(x, y) \sum_{i=1}^{N} \phi_i(l) B_{i,k} \end{aligned} \tag{13}$$

The summations over i is constant for the every pixel so it can be precomputed into $(w_0, \ldots, w_k)$ coefficients where each w_k stores three color components:

$$\rho(x, y, l) = w_0(l) + \sum_{k=1}^{M} a_k(x, y) w_k(l) \tag{14}$$

that is the general formulation of Eq. (2) just discussed. Concluding, in our formulation the light-dependent weights are:

$$\begin{aligned} w_0(l) &= \sum_{i=1}^{N} \phi_i(l) \mu_i \\ w_k(l) &= \sum_{i=1}^{N} \phi_i(l) B_{i,k} \end{aligned} \tag{15}$$

4.1 Luminance-chrominance decoupling

Since PCA minimizes the global reconstruction error in a least square sense, and luminance has usually a much larger variation than chrominance, in general the luminance is better reconstructed than chrominance.

It is possible to control the luminance vs chrominance error introduced by the compression operation, by choosing a suitable color space and applying the compression differently for each component. In particular, we propose to use YCbCr color space and to apply the previous algorithm using a different number of coefficients for the luminance and for the chrominance component:

$$\begin{aligned} \rho_Y(x, y) &\simeq \mu_Y + \sum_{k=1}^{M_Y} a_k(x, y) B_{Y,k} \\ \rho_{Cb}(x, y) &\simeq \mu_{Cb} + \sum_{k=1}^{M_{Cb}} a_k(x, y) B_{Cb,k} \end{aligned} \tag{16}$$

$$\rho_{Cr}(x, y) \simeq \mu_{Cr} + \sum_{k=1}^{M_{Cr}} a_k(x, y) B_{Cr,k} \tag{17}$$

This approach is similar to [Schuster et al. 2014] but in their case only one coefficient is used for the chroma.

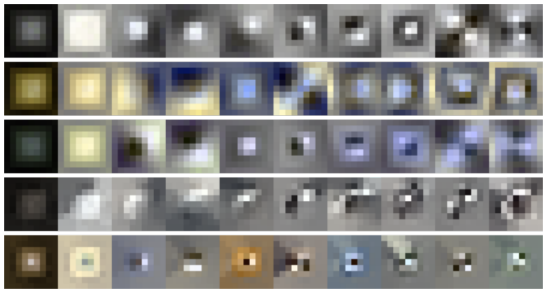

Figure 1: Basis for the BILINEAR encoding of the different datasets. From top to bottom: `tag12783`, `goldcoin`, `bronzecoin`, `horse`, and `panel`. The weights encode chrominance and luminance at the same time. The first column corresponds to the average pattern. The second column encodes great part of the luminance variation. Column 2 and 3 roughly accounts for the surface orientation. Note that the reflectance behaviour of the `horse` dataset (last row), that is a flat surface characterized by great chrominance variations, is well captured.

4.2 Bilinear sampling

The light-dependent weights in (15) can be precomputed only if both the N light directions per pixel and the light direction l are constant along the whole image. The first assumption could be exploited by picking a fixed set of $N' \neq N$ directions and generating a set of re-sampled weights (for example, using again RBF).

Follow this reasoning, we propose a variant of the previous approach, that consists in replacing the gaussian RBF interpolation with a simpler *bilinear interpolation*. Using the concentric projection, where the hemisphere is projected on a pyramid which is then flattened on the square base, we can pick N' light directions in a grid pattern. Then, this pattern is bilinearly interpolated at rendering time to obtain the light-dependent weight $w_k(l)$. An example of the first 10 basis obtained with a 9×9 grid for different datasets (see Results section for a description of the datasets) is shown in Figure 1.

Formally, the resampling is obtained in the following way. Given the original ρ values (R, G, B) at a given pixel we are looking for an array η corresponding to the N' directions. $\rho' = A\eta$ where A is the matrix performing the bilinear interpolation and we want to

minimize $\|\rho - \rho'\|_2$. Generalized Tikhonov regularization [Tikhonov et al. 1995] can be used to find a solution closest to a valid one, and to improve the numerical stability:

$$\min \|A\eta - \rho\|^2 + \tau\|\eta - \eta_0\|^2 \tag{18}$$

The regularization parameter τ was empirically set to 0.1. The initial solution $\eta_0 = R\rho$ corresponds to the one obtained using the RBF interpolation. The closed form solution of this minimization problem can be expressed in matrix form as:

$$\begin{aligned} \eta &= \eta_0 + (A^TA + \tau I)^{-1}(A^T(\rho - Ax_0)) \\ &= R\rho + (A^TA + kI)^{-1}(A^T(\rho - AR\rho)) \\ &= \left(R + (A^TA + \tau I)^{-1}A^T(I - AR)\right)\rho \end{aligned} \tag{19}$$

The same procedure could be applied to different interpolation schemes, to better model the distribution of the light directions.

5 IMPLEMENTATION

We describe here the implementation of the proposed approaches and of the other algorithms implemented for comparison purposes (see the Results section). Since the methods considered shares the general formulation described, most of the code can be easily reused for the different implementations.

The RTI methods can use different basis, such as hemispherical harmonics, bi-quadratic polynomials, etc., and to treat the color components in different ways. For example, a given method can be applied on each color component separately or to all components modulating a base color. Additionally, different color spaces can be considered (e.g. RGB or YCbCr). To limit the combinatorial explosion of these alternative implementations, we restricted us to implement only the most interesting ones. In particular, the following methods have been implemented:

- Polynomial texture maps (LPTM and PTM version as in [Malzbender et al. 2001]).
- Hemi-Spherical harmonics (HSH) ([Zhang and S Drew 2014]).
- Our proposed RBF interpolation plus PCA-based compression method (indicated with RBF in the following).
- Our proposed bilinear interpolation plus PCA-based compression method (indicated with BILINEAR in the following).
- The $\alpha\beta$-JPEG compression method proposed by Schuster al.[Schuster et al. 2014] (denoted with YHSH).

The number of coefficients used is appended to the end of the method name. Hence, RBF18 stands for our proposed RBF+PCA method with 18 coefficients. The same for our BILINEAR method. The other methods implemented have a fixed number of coefficients: 9 for LPTM, 18 for PTM, 27 for HSH and 11 for YHSH.

In order to better control the luminance vs chrominance error, as described in Section 4.1, we implemented also a YCCxyy method, which works in YCbCr color space. YCCxyy means that we apply the proposed RBF+PCA approach with x coefficients for the Y component, and with y coefficients for the Cb and the Cr components, respectively.

Concerning the standard LPTM, we evaluate the base color in the following way: For each pixel, luminance is computed $Y = 0.2126R + 0.7152G + 0.0722B$, then a scaling is applied so that the brightest value is set to 1.0. Finally, a least square minimization is performed to pick the chroma values that result in the minimum global error. In YHSH, following the original paper, the chroma computation is defined as the "average across all the SH coefficients" after a conversion into the YCbCr color space.

For all the methods, the following processing steps are performed:

- Collect all pixels of all images.
- BILINEAR only: for each pixel the light directions are resampled in a 9×9 grid.
- For RBFx, BILINEARx and YCCxyy: perform PCA and compute the basis.
- For LPTM, PTM, HSH, LHSH: compute the coefficients with a matrix multiplication
- Coefficients normalization and quantization.
- Each groups of 3 coefficients is packed into an image.

The coefficients are normalized $c_{norm} = c/scale + bias$ where $scale$ is the difference between the maximum and the minimum value of the coefficients and $bias$ is set to $-minimum/scale$.

The PCA basis is stored as $3Nk$ (8 bits) values in a binary uncompressed file. The final size is few kilobytes. All the parameters, such as the light directions, the number of PCA components, scale, bias, etc. are saved in JSON format, again for a few kB. A direct consequence of this approach is that the format is independent of the algorithm used for the basis computation.

The *real-time rendering algorithm* is basically the same for all the methods: the light dependent weights are computed in Javascript and then setup in the rendering shader. The rendering shader (a pixel shader) simply performs the weighted sum of Eq. (2). Also the color space conversion, if required, is implemented in the shader. Online rendering benefits also from the progressive character of PCA. As the coefficient images are received, the quality increase. Multi-resolution approaches can be taken into account to further improve the performace.

6 RESULTS

The proposed approach and the other algorithms implemented have been tested on five datasets (see Figure 2):

- one gold (named `goldcoin`) and one bronze coin (named `bronzecoin`), from the coins' collection of Palazzo Blu, Pisa, http://vcg.isti.cnr.it/PalazzoBlu;
- a lead tag (named `tag12783`) with commercial inscriptions used in Roman laundry (I-III sec.), part of the Tesserarum Sisciae Sylloge, Zagreb https://tss.amz.hr;
- an oil painting by Giovanni Fattori (named `horse`);
- a bronze panel made by Lorenzo Ghiberti (named `panel`) for the Paradise Door of the Baptistery in Florence.

The first three datasets (`goldcoin`, `bronzecoin` and `tag12783`) have been captured using a dome with 116 light directions [Palma et al. 2012], while the painting and the `panel` have been captured using an hand-held light and 50 lighting directions [Giachetti et al. 2017]. The temporary project web page, with all the experiments and the source code, is available at **http://vcg.isti.cnr.it/relight**.

The images for the test have been cropped to 512×512 pixels in order to better appreciate the details (while `bronzecoin` was resized), and all the methods (PTM, HSH, RBF, BILINEAR, YCC) have been processed using the same JPEG parameters: quality 95, no chroma subsampling. Browser compatibility limits the format

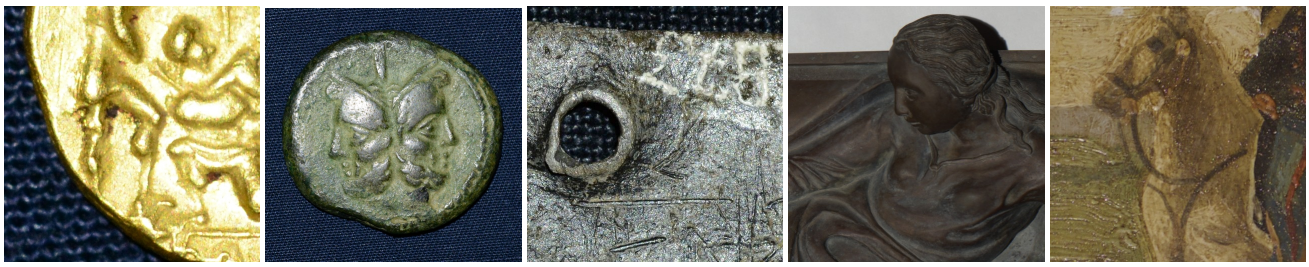

Figure 2: Datasets used for the experimental results (one exemplar for each dataset). From right to left: `goldcoin`, `bronzecoin`, `tag12783`, `panel`, `horse`.

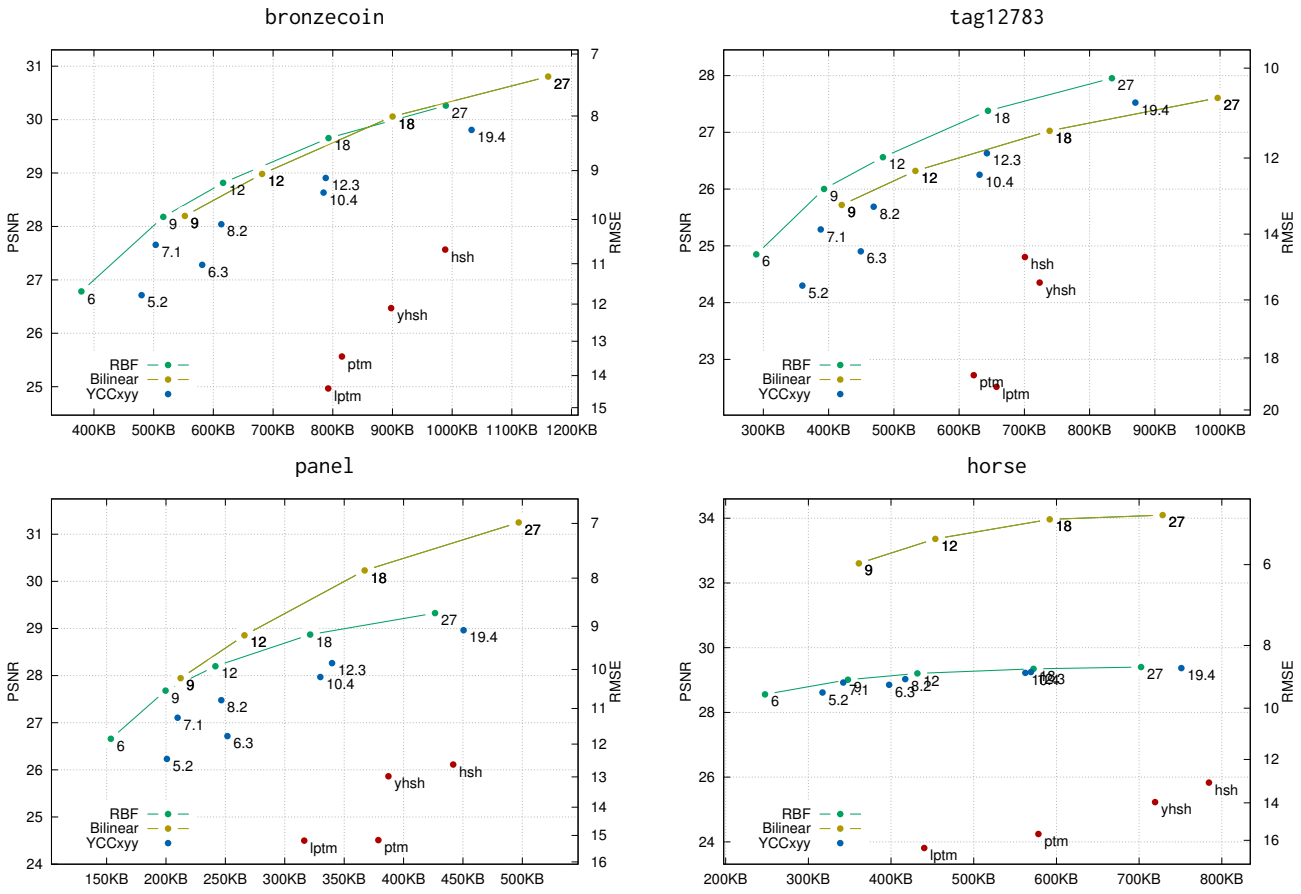

Figure 3: Image quality (PSNR,RMSE) vs size (kB) for different representations and different datasets. For the YCC methods the number X.Y reported in the graphs, indicate the number of luminance and chroma coefficients, respectively.

choice to JPEG and PNG, and the first was preferred in term of compression ratio, for online usage. The standard *libJPEG* library is used for this purpose.

Chroma subsampling can improve the performance for some approaches, but for consistency we did not applied it. Some tests in this direction give us similar qualitative results. For all the methods, the same quantization of the coefficients has been applied, as described in Section 5. Processing a single dataset is in the range of a few seconds, on a consumer PC, for all the methods.

We assess the image quality by comparing the original photographs acquired with the rendering produced under the same light directions. Results are reported in Figure 3 and 4 using Peak Signal-to-Noise Ratio (PSNR) as well as the Structural Similarity Index (SSIM) Wang2004 metrics. We obtain a very good decrease in

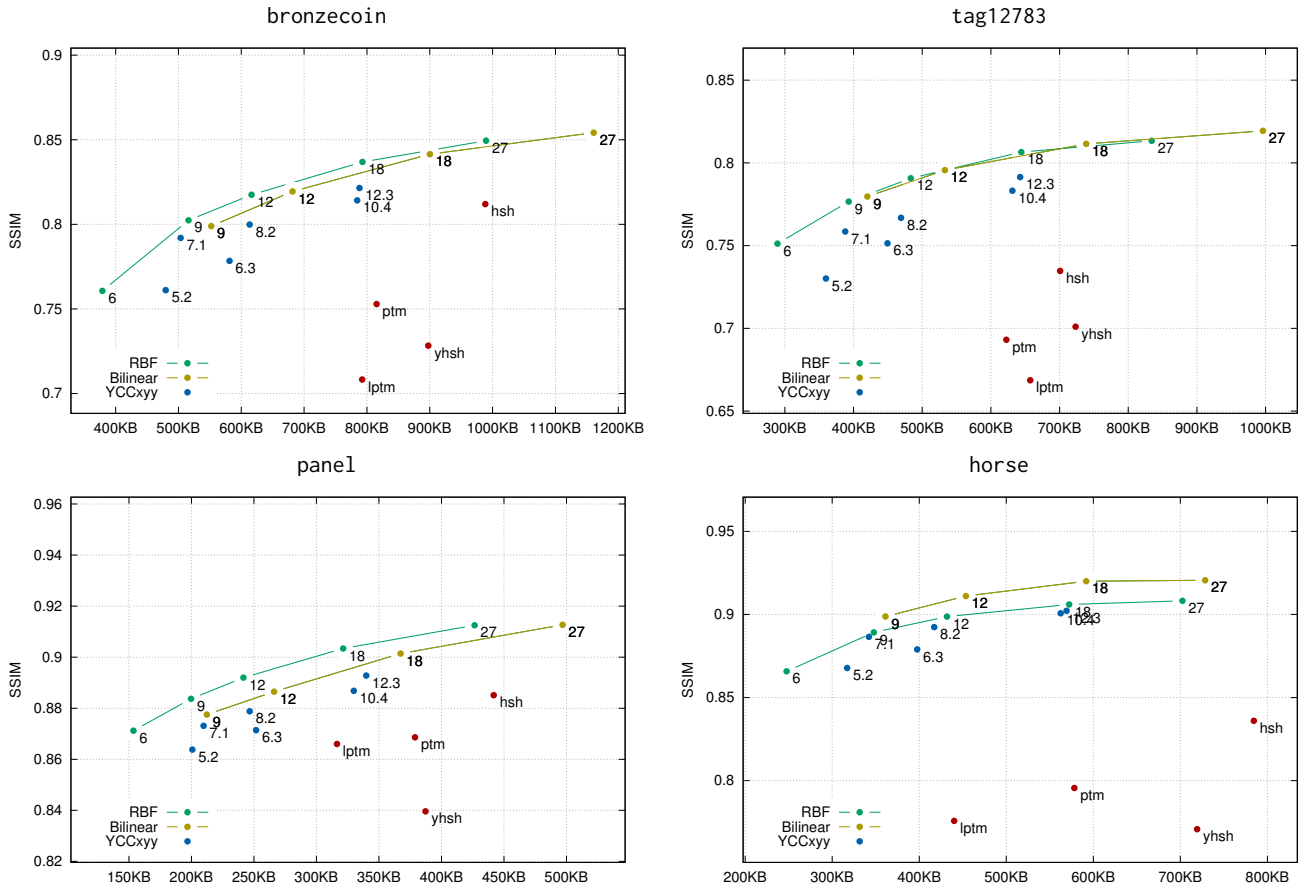

Figure 4: Image quality (SSIM) vs size (kB) for different representations and different dataset. For the YCC methods the number X.Y reported in the graphs, indicate the number of luminance and chroma coefficients, respectively.

size (30-50%) for the same error (or much higher fidelity, also +3-4 dB, for the same size).

For the datasets `horse` and `panel`, the BILINEAR method provided a better reconstruction than RBF. This is due to the number of light directions (50) which is actually smaller than the number of samples used in the bilinear method (81). Additionally, the uneven spacing of the lights forced the σ parameter in the RBF interpolation to be 0.2, higher than the 0.125 value used for the other datasets. This results in a poor interpolation. In general, the resolution of the 9×9 grid used in the BILINEAR method could be increased to slightly improve PSNR with a minimal increase in total size.

Another example of the high quality of the results obtained is shown for the `horse` RTI image, in particular when the light direction is perpendicular ($l_x = l_y = 0$) (see Figure 5). As it is possible to note, YCC711 obtain a visual quality greater than HSH and still requires 40-50% of the disk space.

LPTM and YHSH, allows to trade quality for smaller size. Note that, in some cases (e.g. `tag12783`), it can happen that they perform worse than PTM and HSH, both in term of size and PSNR. This depends on how the JPEG images encode the chromatic and luminance content of the object.

In some cases the average reconstruction error is good but some chromatic visual artifacts remain. For instance, in Figure 6, we marked by red arrows regions where RBF and BILINEAR methods fail to accurately reconstruct the light maroon spot even when faithfully display the reflective property of the gold. As explained in Section 4.1, the reason is that PCA works on luminance and chroma at the same time, but luminance impact prevails in terms of reconstruction error. The final results is that chromatic variations of the same regions can be lost. Luminance and chrominance are treated separately in YCC1233 (in the same figure); it results in a lower PSNR, but accurate chroma reproduction. In the other datasets this problem has not been encountered in the RBF and BILINEAR methods when 27 coefficients (as many as in HSH) are used.

In summary, all the variants of the proposed approach reach a quality/size ratio substantially higher than standard techniques such as PTM or HSH. In general, RBFx ones have the highest quality/size ratio. BILINEAR interpolation is better suited for datasets with a smaller number of lights directions. The use of YCCxyy adds a degree of freedom other than the number of coefficient planes:

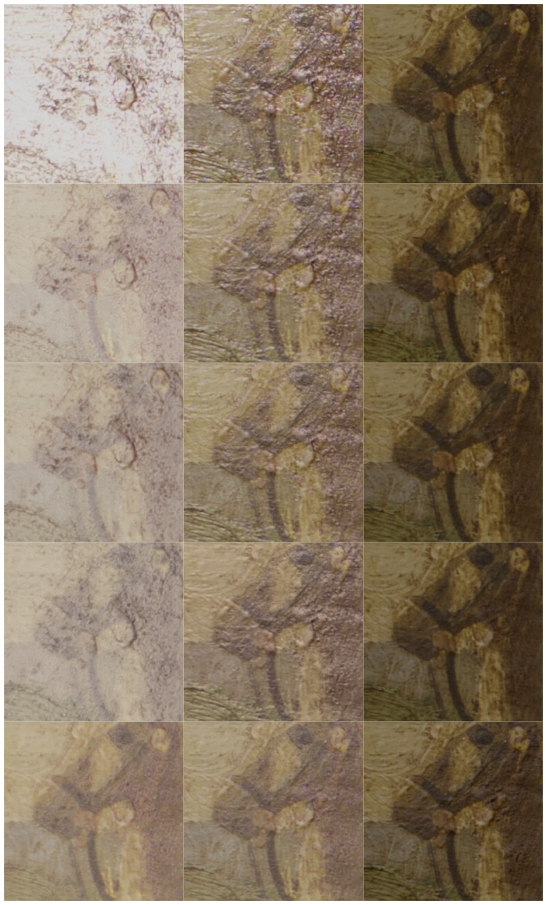

Figure 5: Rows from top to bottom: original images, BILINEAR27, RBF27, YCC711, HSH. The original images shown are the number 0 (column 1), 17 (column 2) and 18 (column 3) of the dataset. Note that the image number 0 has a very strong specular highlight. YCC711 is able to catch this behavior better than the HSH, even if its size is about 50% of the HSH.

the number of coefficients devoted to chrominace vs luminance. This permits to obtain a better chroma fidelity when necessary.

For an in depth visual comparison of the different methods implemented, we remind to the project web page: **http://vcg.isti.cnr.it/relight**.

7 CONCLUSIONS

In this paper we have proposed a compact and efficient web-friendly representation for the RTI images. We have demonstrated that an ad-hoc RTI basis generation, through PCA, allows for low reconstruction error and good compression ratio without increasing the rendering computational cost. The proposed framework allows great control in terms of quality vs size, chromatic fidelity, and high fidelity of the relighted images as the recent interpolation RTI methods. We demonstrated also that the proposed bilinear interpolation variant is a valid alternative to Gaussian RBF, with the advantage of providing good results when the light directions are sampled non-uniformly. It is also possible to extend it to support per-pixel lighting, even if we have not explored this feature in this paper. The code is available as an open source library plus the corresponding tools at the following address: **https://github.com/cnr-isti-vclab/relight** and supports common multiresolution formats such as zoomify and deepzoom.

Some preliminary work to improve the performance by segmenting the images into different materials and applying local PCA has given promising results. Finally, different error metrics for basis creation can be easily explored, taking advantage of the fact that the format and the rendering algorithm needs to specify the basis only.

ACKNOWLEDGMENTS

The research leading to these results has been partially funded by the H2020 Programme under grant agreement no. 654119 (EC "PARTHENOS" project) and by the European Open Science Cloud for Research pilot project (EOSCpilot, EC, DG Research & Innovation under the contract no. 739563).

We wish to thank Andrea Giachetti of University of Verona for the painting dataset, Marion Lamé and the Archaeological Museum in Zagreb for the `tag12783` dataset - funded by the European Association for Digital Humanities - within the digital edition *Tesserarum Sisciae Sylloge* (https://tss.isti.cnr.it), and Gianpaolo Palma and the Palazzo Blu Museum who provided the `coin10` and the `coin16` datasets (http://vcg.isti.cnr.it/PalazzoBlu). Finally, we wish to thank the Cultural Heritage Imaging foundation for providing the reference source code of the *RTIBuilder* and *RTIViewer*.

REFERENCES

Julie Dorsey, Holly Rushmeier, and Franois Sillion. 2008. *Digital Modeling of Material Appearance.* Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Mark S. Drew, Yacov Hel-Or, Tom Malzbender, and Nasim Hajari. 2012. Robust estimation of surface properties and interpolation of shadow/specularity components. *Image and Vision Computing* 30, 4 (2012), 317 – 331.

J. Filip and M. Haindl. 2004. Non-linear reflectance model for bidirectional texture function synthesis. In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, Vol. 1. 80–83 Vol.1.

J. Filip and M. Haindl. 2005. Efficient Image Based Bidirectional Texture Function Model. In *Texture 2005: Proceedings of 4th Internatinal Workshop on Texture Analysis and Synthesis*, M. Chantler and O. Drbohlav (Eds.). Heriot-Watt University, Edinburgh, 7–12.

J. Filip and M. Haindl. 2009. Bidirectional Texture Function Modeling: A State of the Art Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31, 11 (Nov 2009), 1921–1940.

Andrea Giachetti, Irina Mihaela Ciortan, Claudia Daffara, Ruggero Pintus, and Enrico Gobbetti. 2017. Multispectral RTI Analysis of Heterogeneous Artworks. In *Eurographics Workshop on Graphics and Cultural Heritage*, Tobias Schreck, Tim Weyrich, Robert Sablatnig, and Benjamin Stular (Eds.). The Eurographics Association. https://doi.org/10.2312/gch.20171288

A. Giachetti, C. Daffara, C. Reghelin, E. Gobbetti, and Pintus R. 2015. Light calibration and quality assessment methods for Reflectance Transformation Imaging applied to artworks' analysis. *Proc.SPIE* 9527, 9527 – 9527 – 10.

Ø. Hammer, S. Bengtson, T. Malzbender, and D. Gelb. 2002. Imaging fossils using reflectance transformation and interactive manipulation of virtual light sources. *Palaeontologia Electronica* 5, 1 (2002), 9. http://palaeo-electronica.org/2002_1/fossils/issue1_02.htm

Pun-Mo Ho, Tien-Tsin Wong, and Chi-Sing Leung. 2005. Compressing the illumination-adjustable images with principal component analysis. *IEEE Transactions on Circuits and Systems for Video Technology* 15, 3 (March 2005), 355–364.

Melissa L Koudelka, Sebastian Magda, Peter N Belhumeur, and David J Kriegman. 2003. Acquisition, compression, and synthesis of bidirectional texture functions. In *3rd International Workshop on Texture Analysis and Synthesis (Texture 2003)*. 59–64.

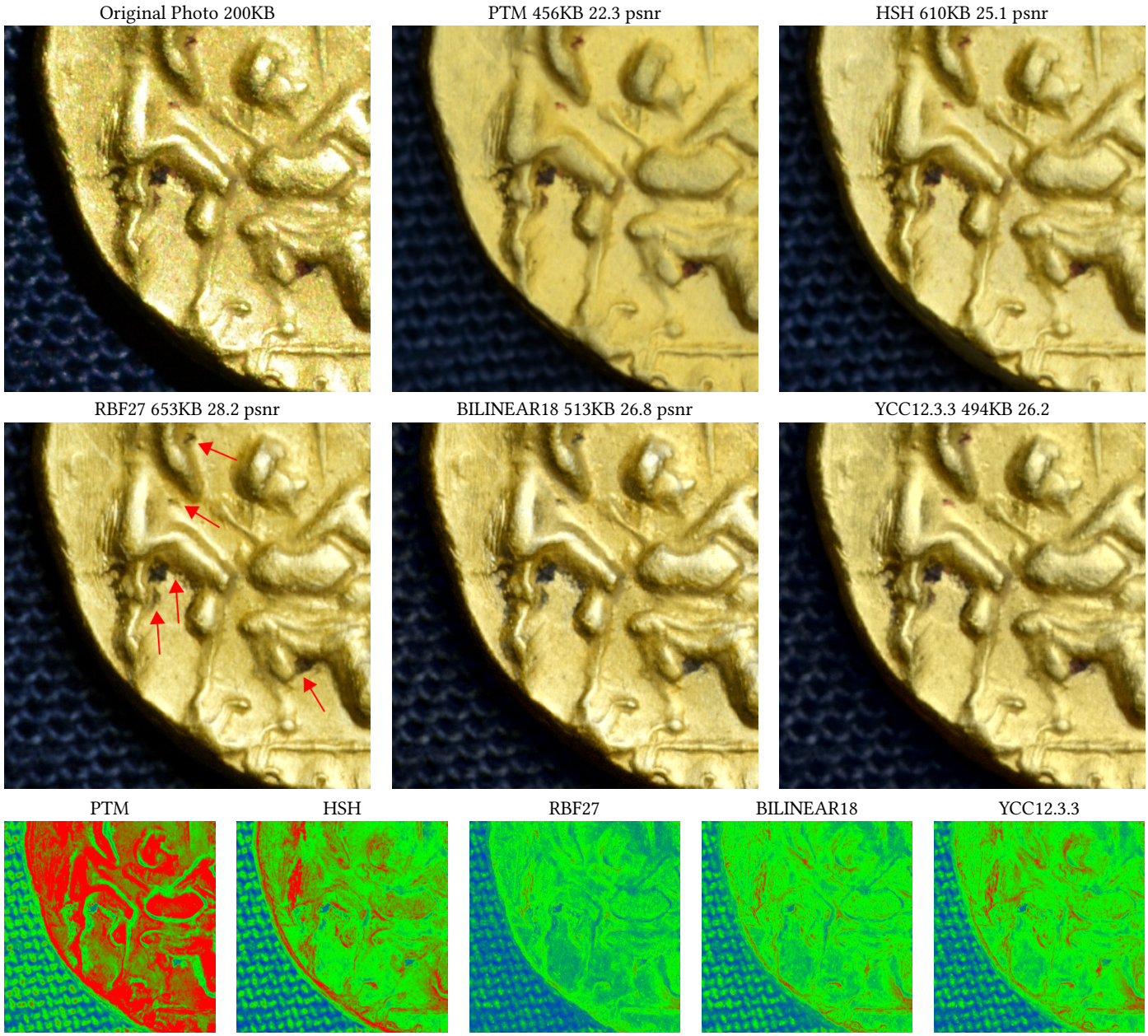

Figure 6: Visual comparisons. Original photograph (top-left). PTM, HSH, RBF27 and BILINEAR18 rendered from the same light direction. (Bottom) Color-coded RMSE (ranging from 0, blue, to 25, red). See more results at the project web page: http://vcg.isti.cnr.it/relight.

P. M. Lam, C. S. Leung, and T. T. Wong. 2012. Noise-resistant hemispherical basis for image-based relighting. *IET Image Processing* 6, 1 (Feb 2012), 72–86.

Marion Lamé. 2015. Primary Sources of Information, Digitization Processes and Dispositive Analysis. In *Proceedings of the Third AIUCD Annual Conference on Humanities and Their Methods in the Digital Ecosystem (AIUCD '14)*. ACM, New York, NY, USA, Article 18, 18:1–18:5 pages.

Tom Malzbender, Dan Gelb, and Hans Wolters. 2001. Polynomial texture maps. In *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. ACM, 519–528. https://doi.org/10.1145/383259.383320

Jan Meseth, Gero Müller, and Reinhard Klein. 2003. Preserving Realism in Real-Time Rendering of Bidirectional Texture Functions. In *OpenSG Symposium 2003*. Eurographics Association, Switzerland, 89–96.

Mark Mudge, Tom Malzbender, Alan Chalmers, Roberto Scopigno, James Davis, Oliver Wang, Prabath Gunawardane, Michael Ashley, Martin Doerr, Alberto Proenca, and Joao Barbosa. 2008. Image-Based Empirical Information Acquisition, Scientific Reliability, and Long-Term Digital Preservation for the Natural Sciences and Cultural Heritage. In *Eurographics 2008 - Tutorials*, Maria Roussou and Jason Leigh (Eds.). Eurographics Association, Crete, Greece. http://www.eg.org/EG/DL/conf/EG2008/tutorials/T2.pdf

Mark Mudge, Jean-Pierre Voutaz, Carla Schroer, and Marlin Lum. 2005. Reflection Transformation Imaging and Virtual Representations of Coins from the Hospice of the Grand St. Bernard. In *The 6th International Symposium on Virtual Reality,*

Archaeology and Cultural Heritage, Mark Mudge, Nick Ryan, and Roberto Scopigno (Eds.). Eurographics Association, Pisa, Italy, 29–39. https://doi.org/10.2312/VAST/VAST05/029-039

Gero Müller, Jan Meseth, and Reinhard Klein. 2003. Compression and Real-Time Rendering of Measured BTFs Using Local PCA. In *Vision, Modeling and Visualisation 2003*, T. Ertl, B. Girod, G. Greiner, H. Niemann, H.-P. Seidel, E. Steinbach, and R. Westermann (Eds.). Akademische Verlagsgesellschaft Aka GmbH, Berlin, 271–280.

Joseph Padfield, David Saunders, and Tom Malzbender. 2005. Polynomial texture mapping: a new tool for examining the surface of paintings. In *ICOM Committee for conservation triennial meeting*. Eurographics Association, Netherlands, 504–510.

Gianpaolo Palma, Eliana Siotto, Marc Proesmans, Monica Baldassarri, Clara Baracchini, Sabrina Batino, and Roberto Scopigno. 2012. Telling The Story Of Ancient Coins By Means Of Interactive RTI Images Visualization. In *CAA 2012 Conference Proceeding*. Pallas Publications - Amsterdam University Press (AUP), 177–185. http://vcg.isti.cnr.it/Publications/2012/PSPBBBS12

Ruggero Pintus, Andrea Giachetti, Giovanni Pintore, and Enrico Gobbetti. 2017. Guided Robust Matte-Model Fitting for Accelerating Multi-light Reflectance Processing Techniques. In *Proc. British Machine Vision Conference*.

Mirko Sattler, Ralf Sarlette, and Reinhard Klein. 2003. Efficient and Realistic Visualization of Cloth. In *Proceedings of the 14th Eurographics Workshop on Rendering (EGRW '03)*. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 167–177.

C. Schuster, B. Zhang, R. Vaish, P. Gomes, J. Thomas, and J. Davis. 2014. RTI compression for mobile devices. In *Proceedings of the 6th International Conference on Information Technology and Multimedia*. 368–373. https://doi.org/10.1109/ICIMU.2014.7066661

Christopher Schwartz, Roland Ruiters, Michael Weinmann, and Reinhard Klein. 2013. WebGL-based Streaming and Presentation of Objects with Bidirectional Texture Functions. *Journal on Computing and Cultural Heritage (JOCCH)* 6, 3 (July 2013), 11:1–11:21. https://doi.org/10.1145/2499931.2499932

A.N. Tikhonov, A. Goncharsky, V.V. Stepanov, and A.G. Yagola. 1995. *Numerical Methods for the Solution of Ill-Posed Problems*. Springer.

Mingjing Zhang and Mark S Drew. 2014. Efficient robust image interpolation and surface properties using polynomial texture mapping. *EURASIP Journal on Image and Video Processing* 2014, 1 (01 May 2014), 25. https://doi.org/10.1186/1687-5281-2014-25