https://doi.org/10.2352/issn.2168-3204.2023.20.1.15

A benchmark dataset and evaluation for best light configuration in Reflectance Transformation Imaging

Ramamoorthy Luxman[1];, Hermine Chatoux[1] ; Gaetan Le Goic[1] ; Jon Yngve Hardeberg[2] ; Franck Marzani[1] ; Alamin Mansouri[1]
[1]ImViA Laboratory, Université de Bourgogne, Dijon, France. Université de Bourgogne
[2]Department of Computer Science, Norwegian University of Science and Technology (NTNU), Gjøvik, Norway.

Abstract

Reflectance Transformation Imaging (RTI) is a non-invasive technique that enables the analysis of materials. Recent advancements in this technology, along with the availability of software for surface analysis through relighting, have improved the restoration and conservation of cultural heritage objects. However, there is a lack of appropriate benchmark data and reference light configurations, which makes it difficult to quantitatively compare and evaluate RTI data acquisitions. To address this, we have developed a dataset that can be used to assess the effectiveness of different surface light configurations for RTI acquisition. Additionally, we introduce methods to derive an ideal reference light configuration for a surface from its dense RTI acquisition. This dataset provides a standardized set of dense RTI acquisitions, accompanied by their corresponding reference light configurations that were obtained using our methods. This dataset can help researchers and developers to compare the performance of their approaches in solving the "Next Best Light Position" problem in RTI acquisition, which can ultimately improve the accuracy and efficiency of RTI acquisition and broaden its applicability in various fields.

Motivation

RTI [1] [2] has emerged as a simple yet powerful tool for visual analysis and surface characterization particularly in the field of cultural heritage. It has been extensively researched in post-acquisition steps such as modeling and visualization, leading to numerous advancements in these areas. However, the acquisition process itself has not been studied extensively, and there are no sophisticated methods available to ensure quality RTI acquisition that is adaptive to the object being captured. The best light positions for RTI will depend on the specific object or scene being imaged and the desired outcome. The light sources should be positioned to create a range of highlights and shadows on the object, allowing for the capture of a wide range of surface details. The accuracy of RTI heavily depends on the quality and number of images captured. The positioning and direction of the light sources play a significant role in the quality of the resulting images. However, finding the optimal light positions for RTI requires a considerable amount of expertise and trial-and-error experimentation. This reduces the efficiency of acquisition greatly. To address this problem, we have created a benchmark dataset and evaluation for the best light positions in RTI [3].

Problem

To investigate the impact of light position selections on the performance and quality of Reflectance Transformation Imaging (RTI), we conducted an evaluation of RTI data quality surfaces using different light configurations. We used brushed metal with a dent as a case study and performed very dense RTI acquisition (1000 homogeneously distributed light positions) which were considered as ground truth. We then performed sparse acquisitions ith homogeneously distributed light positions incrementing from 40 number of positions to 55, 80 and so on. We then created DMD model for each acquisition. We analyzed the degradation of RTI quality in these sparse acquisitions through comparison of the normal maps and reconstructed image from the dense DMD and each of the subsampled ones. Figure. 1 shows the degradation in the quality of RTI with different light configurations. The top row displays the positions of lights projected on a 2D plane. The middle row displays the reconstructed images from the respective sparse acquisitions (relighted from an elevation of 45° and an azimuth of 40°). The bottom row shows maps of the difference between the dihedral angles of ground truth normal and the normal obtained from corresponding sparse RTI acquisitions. It can be observed the light configuration does affect the quality of RTI data. Hence finding the best light configuration for performing RTI acquisition is a open problem that requires to be solved.

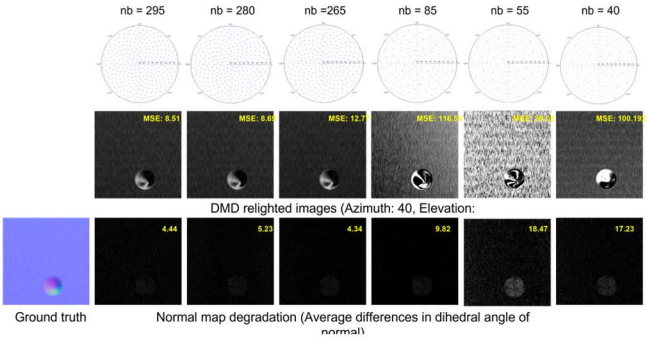

Figure 1 Degradation in the quality of RTI with different light configurations

Our objectives in this work are twofold:

1. Acquire RTI data of various objects and create a dataset.
2. Introduce an approach to deduce reference ideal set of light directions for each surface from dense acquisition of them.

The deduced reference acquisitions containing the ideal set of light directions should have the following characteristics:

1. The estimated best reference light directions must be a unique set of light positions and must have no or least number of redundant measurements in it. The acquisition must capture the luminance behavior of all points on the surface as close as possible to a dense acquisition, but with the fewest number of light directions.

2.The gradients of the signals in a reference NBLP acquisition should be as uniform as possible, that is, the light position is densely distributed along the directions where the luminance of a surface point changes rapidly, and sparsely distributed along the directions where the luminance change is small.

We built this benchmark dataset and evaluation framework with an aim to provide a standardized and objective approach for evaluating the performance of Next Best Light Position algorithms for RTI. This will facilitate the development of new algorithms and techniques that can automatically determine the optimal light positions for RTI, thereby making the process more efficient and accessible to a wider range of users.

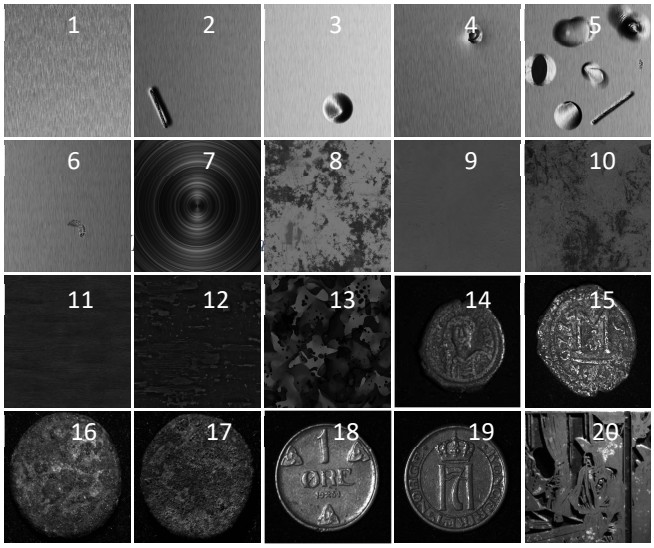

Figure 2 (The reference NBLP points on each ring are shown together)

Approach

Dataset creation: The NBLP-RTI dataset was created to evaluate NBLP methods for estimating the optimal light configuration for a surface under standard RTI imaging set-up, where the camera is fixed orthogonally to the surface. The dataset comprises high-resolution images of a range of objects as shown in Figure. 3 captured under a variety of lighting conditions. This dataset focuses on anisotropic, non-Lambertian, semi-glossy, non-Lambertian diffuse surfaces. There are some publicly available datasets for RTI and RTI-like techniques **[4] [5] [6]**. But none of them are suitable for evaluating methods for detecting ideal lighting directions. We therefore created a dataset of RTI acquisitions of virtual and real surfaces, accompanied by very dense RTI acquisitions and ground truth 3D data. For real surfaces, RTI acquisitions were made using a carefully calibrated mechanized RTI dome setup built in-house, and 3D shapes were acquired using a structured light scanner. This dataset contains 20 surfaces, as shown in Figure. 3. Surfaces 1 to 13 are virtual surfaces created in Blender software using physically based rendering methods and surfaces 14 to 20 are real objects. The physical dimensions of the surfaces range from 2 cm to 4 cm wide and 3 mm thick. With regards to BRDF, the dataset covers anisotropic surfaces (brushed metals, 1 to 7), isotropic metallic surfaces (8,9), Lambertian surfaces with strong details and nearly homogenous specular lobes (11,12,13) and non-Lambertian diffuse surfaces with strong specular spikes (14 to 20).

Creating a reference for an ideal acquisition by using information from a dense acquisition: In our approach, we derive reference ideal light directions for the RTI acquisition of a surface from a very dense acquisition of the surface. This may seem like an indirect approach, but it is justified because there are no ground truths for RTI acquisitions. We present two approaches for deducing the reference light positions based on specific case scenarios:

1. ***Ring acquisitions (azimuth only space)*:** For characterizing the reflectance of a surface, the azimuth space is often more important than the elevation space. Many analysis techniques use ring acquisitions to create maps based on directional slopes and curvature. On the other hand the reflectance behavior of the surface in the elevation space can be interpolated using a second-order polynomial with just a few sample measurements.

To efficiently capture the non-linear nature of reflectance in the azimuthal space, we introduce a method for finding best azimuth angles for a surface from its dense RTI acquisitions. The goal of the sampling strategy is to reduce a dense acquisition of surface reflectance signals to a sparse representation without losing much information. The dense acquisition comprises surface points illuminated from different directions in the azimuth space from 0° to 360° . Firstly the surface points are classified as diffuse, semi-specular, or specular based on the maximum gradient observed in its reflectance signals. If the maximum gradient of any signal is greater than 10, the surface is considered to have some specular or semi-specular points. Otherwise, the surface is considered diffuse. To cluster the signals, we use the K-means temporal signal clustering technique proposed by [7]. For diffuse reflectance signals, we simply take 8 to 10 evenly spaced light positions as suitable for the RTI acquisition. For specular and semi-specular signals, we apply an approach as described below to identify their pertinent light directions.

To achieve this, we traverse from 0° to 360° in the dense acquisition and sample a point if it satisfies any of the following three criteria:

- The current sampling point is spaced higher than a threshold distance (arc length) to the previous sampled point. This criterion ensures that we sample points at regular intervals, rather than sampling points too close together.
- The gradient of the signal at a point is higher than a threshold. This criterion ensures that we sample points where there is a significant change in the signal, as these points are likely to be more informative than points with little or no change.

- The difference between the current signal value and the previously sampled signal value is higher than the threshold. This criterion ensures that we sample points where there is a significant change in the signal, even if the gradient at that point is not very high.

For calculating the threshold, we use two different ways depending on the nature of reflectance. If the reflectance is specular in nature, we use half the highest gradient observed in the dense reflectance signal. If the reflectance is diffuse, we use twice the highest gradient value observed in the dense reflectance signal. This is done to maintain uniform gradient in the decimated reflectance signal. This strategy is illustrated in Figure. 4

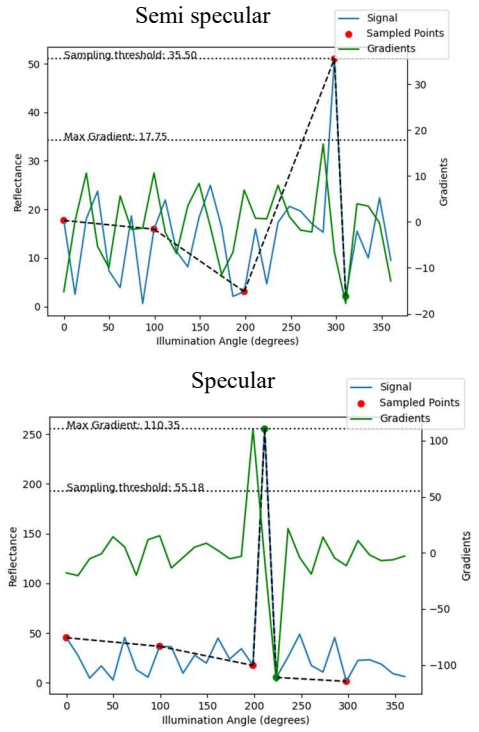

Figure 4 Examples of reflectance signals measured of two different points belonging to the same surface.

Mathematically, we can represent the decimation process as follows:

Let S be the set of all points in the dense acquisition, and $s_i \in S$ be an individual point. Let be the distance between points s_i ands_i, $g(s_i)$ be the gradient of the signal at point s_i. Let T_d be the threshold distance for sampling, T_v be the threshold gradient or difference in signal value for sampling. The set of sampled points S' can be defined as:

$$S' = \{s_i \in S | d(s_i, s_{i-1}) > T_d \vee g(s_i) > T_v \vee |v(s_i) - v(s_{i-1})| > T_v\}$$

where, $T_v = \frac{1}{2} max(g(s_i))$ if the reflectance is specular and $T_v = 2 \times max(g(s_i))$ if the reflectance is semi specular. It is computationally challenging to sample all the signals (surface points) simultaneously. To make the process efficient, we First gather all the dense acquisition of surface reflectance signals and put them into a single matrix called the acquisition data matrix M. Then to perform the proposed sampling method efficiently, we use matrix column operations instead of looping through each signal and acquisition point individually. This allows us to apply the sampling method to all signals simultaneously, which is much faster and more efficient than processing each signal separately.

2. ***Azimuth-Elevation space*:**

Adding the elevation dimension to the problem greatly increases its complexity. While gradient-based sampling is an effective technique for analyzing 1D signal decimation, it becomes much more challenging when applied to azimuth-elevation data. This is because the non-continuous dense signal contains data that are not arranged in a regular grid. For performing gradient descent, computing partial derivatives along the x and y directions by regularizing the grid using any surface fitting methods for millions of signals would be computationally prohibitive.

Decimating multiple signals simultaneously is challenging to preserve the characteristics of each signal while reducing their complexity. Our approach to solving this problem is to represent each signal as a 3D pointcloud and decimate these pointclouds. This approach has the advantage of allowing for more control over the decimation process, and it can result in more accurate and faithful representations of the original signals.

Similar to our approach in the azimuth-only space, the surface points are first classified as diffuse, and specular. However, in azimuth-elevation space, computing the signal gradients by comparing each signal point to its neighbours is computationally very expensive. Instead we identify the specular and semi specular signals by simply comparing the minimum and maximum values observed in a signal. If the difference between the max and min is lower than 20, we classify the signal as diffuse, if it is higher than 20 we classify it semi specular. This is an approximate classification a for the preliminary step of the decimation process. To optimize the light positions, we focus on specular surface points. These light positions optimized for the specular surface points are also expected to work well for diffuse points, as the reflectance of diffuse points is uniform.

In our approach we create a pointcloud for each signal. To create a pointcloud for a signal, we create points in 3D space by combining the x and y coordinates of the light positions with the normalized signal values as z coordinates. We then decimate the pointclouds using the Quadric Error Metrics (QEM) decimation approach **[7] [8]**. This method works by iteratively removing points from the point cloud, starting with the point that has the smallest quadric error when removed till we reach a desired size of the pointcloud. The desired size of the pointcloud corresponds to the desired number of light positions for the acquisition and it can be set to any value

depending on the requirement. For PTM, DMD, any number of points > 50 is generally considered reasonable size of an acquisition.

To compute the quadric error for a given point in the point cloud, a quadric matrix is first constructed to represent the local geometry of the points in the neighborhood of that point. The quadric matrix is constructed using a least squares fit to the points in the neighborhood of the point being considered for removal. This fit is used to estimate the local geometry of the points, and the resulting quadric matrix is used to evaluate the error introduced by removing the point. The quadric error for a point is given by the following equation:

$$E = (x - \hat{x})^T Q (x - \widehat{x})$$

where x is the original position of the point, $\hat{x}$ is the position of the point after it has been removed, and Q is the quadric matrix as shown in Figure 5.

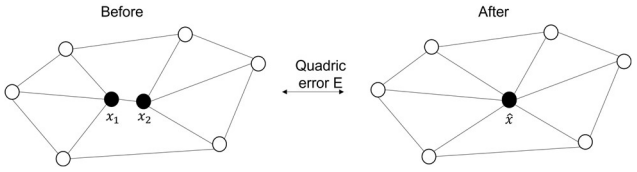

Figure 6Quadric error of a point being removed

The quadric matrix is given by the equation:

$$Q = \sum_{i=1}^{n} (x_i - \bar{x})(x_i - \bar{x})^T$$

where x_i are the positions of the points in the neighborhood of the point being considered for removal, and $\bar{x}$ is the centroid of these points. The QEM decimation algorithm proceeds by iteratively removing the point with the smallest quadric error from the point cloud until the desired number of points is reached. This ensures that the points removed are the ones that introduce the least amount of error, and thus the shape of the point cloud is preserved as much as possible.

Calculating the quadric error for each point in a signal can be computationally expensive, especially when dealing with a large number of such signals. The quadric error minimization (QEM) decimation algorithm has a time complexity of $O(n^2)$ where n is the number of points in the point cloud. This is because the quadric error for each point must be computed individually. However, we can use a matrix kernel to compute the quadric error for all points simultaneously, reducing the time complexity to $O(n)$. To do this, we first compute the quadric matrix for each point using the points in its neighborhood. Then, we subtract the centroid of each neighborhood from each point in that neighborhood to obtain a set of centered points.

Finally, we compute the quadric error for all points simultaneously by taking the dot product of the centered points with the quadric matrix and the transpose of the centered points. This operation is available in standard programming libraries. Using a matrix kernel to compute the quadric error allows us to significantly improve the performance of the QEM decimation algorithm. This makes it possible to decimate large number of sigmals more efficiently, and thus can be applied to derive reference best light positions in the azimuth-elevation space.

Results

We present the results for selected surfaces representing the general behavior of each surface category

Azimuth only space:

Light positions : Figure 6 depicts the reference best light positions obtained using the proposed method for the brushed metal plate (Surface 1), the brushed metal plate with a dent defect (Surface 2), the canvas painting (Surface 13), and the antique coin 1 (surface 14). Surfaces 1, 2, 13 are virtual surfaces and surface 14 is a real surface. It can be observed that the proposed method is able to adaptively derive the reference best light positions to the surface.

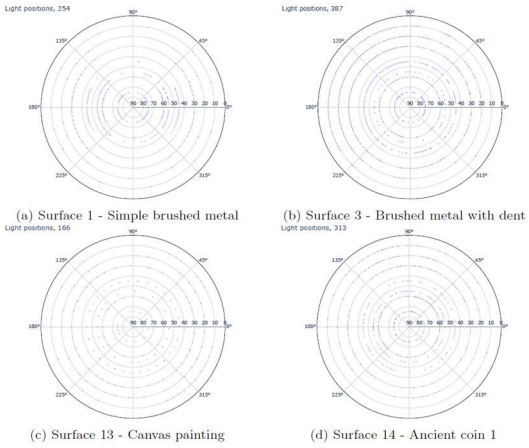

Figure 5 The reference ideal points on each ring are shown together, showing the overall distribution of surface light positions

Since surface 13 is a perfectly diffuse surface, the number of reference best light positions is small and nearly evenly distributed along the rings. For anisotropic surface 1, the method correctly identifies the direction exhibiting anisotropy and samples more points along that direction than other directions to preserve gradient uniformity. Surface 3, Surface 14 contain details showing specular reflections of many random surface points, so the distribution of reference light positions is numerically larger.

Reflectance signal gradients: The proposed method can achieve the goal of maintaining the gradient uniformity of the reflected signal. Figure. 7 compares the gradient at a single surface point (here we have chosen the pixel showing the largest gradient) in the reference best light positions acquisition versus the dense acquisition.

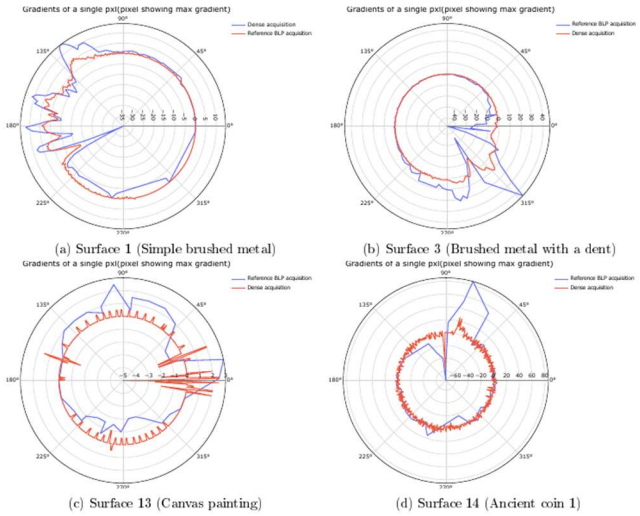

Figure 7Comparison of reflectance signal gradients from dense and reference best light position acquisition of surfaces

Azimuth-elevation space:

Light positions: Figure. 8 shows the reference best light positions obtained using the proposed method for azimuthal-elevation space for the four surfaces.

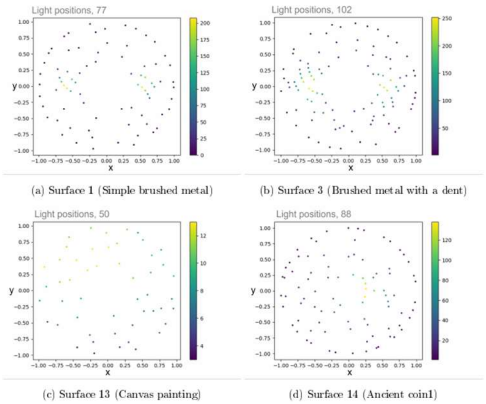

Figure 8 Best light positions obtained for the surfaces from the respective dense acquisition

In this example, the minimum number of light positions is set to 50 for all surfaces. The method adaptively generates a set of light positions that are most effective at capturing the reflectance of the surface. For the perfectly diffuse surface 13, the number of reference light positions is equal to the minimum number set and is evenly distributed along rings. For the anisotropic surface 1, the specular reflections are regular and symmetric, so most specular points have similar directions. Surface 3, which has details around a dent and a brush feature, has many more specular directions, and the method successfully identifies and samples them. Surface 14 has details with specular reflections at many random surface points, so the distribution of reference light positions is concentrated in random regions and evenly spread in most other regions.

In contrast to ring acquisitions (which only have azimuth space), azimuth-elevation space acquisitions have both strong linear (elevation) and strong non-linear (azimuth) behavior. Therefore, using dispersion of gradients as a measure of performance is not appropriate for this case. Instead, we compare the reconstruction errors between the best light positions derived using our method, dense light positions, and sparsely homogeneous light positions. To illustrate the impact of the choice of light position on RTI modeling, we perform a statistical analysis of the reconstruction errors. For that, we consider the following four acquisitions.

1. Uniformly distributed dense acquisition (1000 unique directions)

2. Uniformly distributed sparse acquisition (100 unique directions)

3. Reference best light positions acquisition (number of unique directions adaptive to the surfaces.

4. Test acquisition containing 500 random light positions.

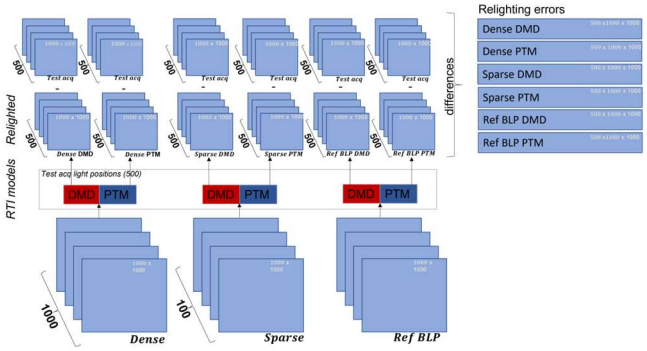

Figure 9 Acquisitions carried out to compute the reconstruction error statistics

Our approach to the statistical analysis of reconstruction errors is illustrated in Figure.9. We perform DMD and PTM RTI modeling on acquisitions with uniform distribution of dense light positions, uniform distribution of sparse light positions, and reference best light positions acquisitions. We then use the fitted model to relight the surface from the test acquisition light positions. The relighted images are compared to the corresponding actual captured images in the test acquisition by calculating the absolute point to point differences. The comparison gives the errors in the surface points reflection reconstruction. We estimate the probability density function and cumulative distribution function **[9]** of the measured errors in both DMD and PTM relighted images.

Figure. 10 shows the PDF and CDF plots of the reconstruction errors of the surfaces.

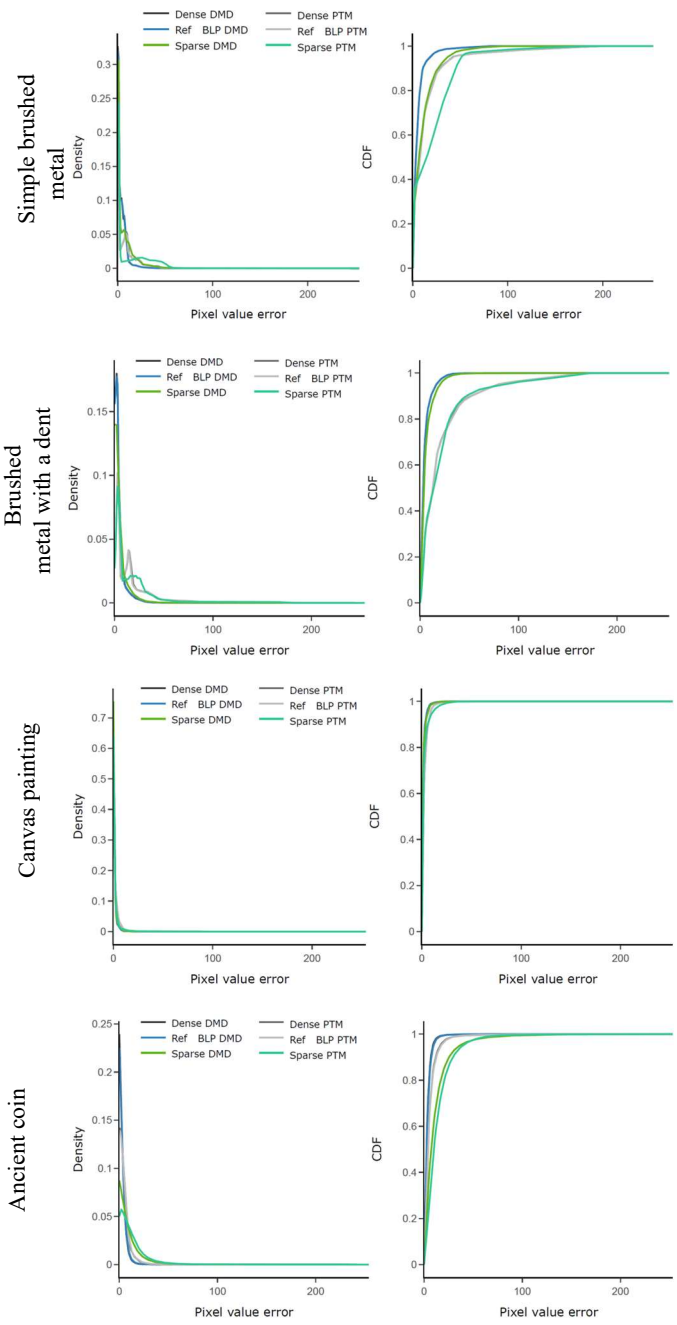

Figure 10 PDF and CDF of the reconstruction errors of Surfaces

Conclusion

In this paper, we present a benchmark dataset containing dense acquisitions and reference ideal acquisitions with best light positions derived using a proposed novel method for both azimuth-only space acquisitions and azimuth elevation space acquisitions. Our results showed that the proposed method was able to effectively generate a set of best light positions that improves the reconstruction of reflectance maps in both acquisition spaces. The contributions of our work include the development of novel methods for generating reference best light positions from dense acquisition and the creation of a benchmark dataset that can be used to evaluate the performance of different NBLP strategies.

References

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

[2] G. Pitard, G. Le goïc, A. Mansouri, H. Favrelière, S.-F. Desage, S. Samper and M. Pillet, "Discrete Modal Decomposition: a new approach for the reflectance modeling and rendering of real surfaces," *Machine Vision and Applications,* vol. 28, pp. 607-621, 15 July 2017.

[3] R. Luxman, M. Nurit, G. Goïc, F. Marzani and A. Mansouri, "Next Best Light Position: A self configuring approach for the Reflectance Transformation Imaging acquisition process," *Electronic Imaging,* vol. 33, pp. 132-1-132-7, 18 January 2021.

[4] H. Aanæs, A. Dahl and K. Steenstrup Pedersen, "Interesting Interest Points," *International Journal of Computer Vision,* vol. 97, pp. 18-35, 2012.

[5] R. Jensen, A. Dahl, G. Vogiatzis, E. Tola and H. Aanaes, "Large Scale Multi-view Stereopsis Evaluation," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2014, p. 406–413.

[6] B. Shi, Z. Wu, Z. Mo, D. Duan, S.-K. Yeung and P. Tan, "A Benchmark Dataset and Evaluation for Non-Lambertian and Uncalibrated Photometric Stereo," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2016, p. 3707–3716.

[7] M. Garland and P. Heckbert, "Surface simplification using quadric error metrics," in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques - SIGGRAPH '97*, 1997.

[8] J. Yu, M. Wei, J. Qin, J. Wu and P.-A. Heng, "Feature-preserving mesh denoising via normal guided quadric error metrics," *Optics and Lasers in Engineering,* vol. 62, pp. 57-68, November 2014.

[9] M. Tamayo, S. Valcárcel Andrés, J. Pons and M, "Applications of reflectance transformation imaging for documentation and surface analysis in conservation," *International Journal of Conservation Science,* vol. 4, p. 535–548, 2013.

[10] A. Grønlund, K. Larsen, A. Mathiasen, J. Nielsen, S. Schneider and M. Song, "Fast exact k-means, k-

medians and Bregman divergence clustering in 1D," 2017.

Ramamoorthy Luxman is a PhD candidate at Universitéde Bourgogne, France. He has M.Sc. in Robotics Engineering.

Hermine Chatoux is an Assistant Professor at the University of Bourgogne since 2020 and a member of ImVia laboratory. Her current research is focused on colour sciences and spectral image processing with a metrological purpose. She is interested in experimentation to analysis and understanding the human visual system/the physical sense.

Gaëtan Le Goïc is an Assistant Professor at the University of Bourgogne and a member of ImVia laboratory. His research activity lies at the intersection of the fields of mechanical engineering (mechanical surface functions, metrology, quality) and the field of imaging and signals (acquisition, analysis and processing of measured data).

Jon Y. HARDEBERG received his PhD degree from Ecole Nationale Supérieure des Télécommunications in Paris, France in 1999. He is now Professor of Colour Imaging at the Colourlab at NTNU – Norwegian University of Science and Technology, Gjøvik, Norway. His current research interests include spectral imaging, image quality, colour management, material appearance, and cultural heritage imaging, and he has co-authored more than 300 publications within the field.

Franck MARZANI obtained his Ph.D. in computer vision and image processing in 1998. He is full professor at the Université de Bourgogne, Dijon, France since 2009. He is currently the head of the ImViA research laboratory (Imaging & Computer Vision). His research interests include acquisition and analysis of images. He has been developing an activity on feature extraction from color and multispectral images for classification purposes. These methodologies have been proposed in the frame of different applications.

Alamin MANSOURI is a full professor at the university of Bourgogne since 2015 and a member of ImViA laboratory where he is co-leading the CORES team. His current research is focused on multimodal imaging for appearance capture and modelling with main applications in Cultural Heritage and Industry.