Self-supervised classification of surfaces using reflectance transformation imaging

Muhammad Arsalan Khawaja[1,2,*], Sony George[2], Franck Marzani[1], Jon Yngve Hardeberg[2] and Alamin Mansouri[1]

[1]ImViA Lab, Université de Bourgogne, 21000 Dijon, France
[2]Colourlab, Department of Computer Science, Norwegian University of Science and Technology, 2815 Gjøvik, Norway

Abstract

Reflectance Transformation Imaging (RTI) is an imaging technique used to analyze objects or surfaces by capturing their appearance under varying illumination directions. This paper proposes two self-supervised learning algorithms to classify surfaces according to their reflectance profiles. The classification problem is addressed using K-means and Self Organizing Map (SOM) neural networks. The proposed methodology is evaluated using both real and synthetic datasets. The primary motivation for our approach is to exploit illumination variation data to enhance surface understanding and detect anomalies. Given the exploratory nature of this task and the lack of ground truth for comparison, a self-supervised method was deemed most suitable. The classification of surfaces using reflectance information has immense applications in fields such as Cultural Heritage (CH) preservation, digitization, and industrial quality control.

Keywords

Material Appearance, Reflectance Transformation Imaging (RTI), Self-Supervised Learning, Reflectance profile, Cultural Heritage (CH)

1. Introduction

Reflectance Transformation Imaging (RTI) is a non-contact, non-destructive, computational imaging technique used to study objects under varying illumination directions. This provides an enhanced visualization experience which is its primary use. RTI was first developed at Hewlett Packard Laboratories (HP Labs) in 2001 [1]. Since then, many new RTI tools (PTMFitter, RTIBuilder and RTIViewer) and methods have been developed to exploit this imaging technique for various applications. RTI has three stages. The first stage is *Acquisition*. It encompasses the acquisition setup and planning algorithms, calibrations, preparation, and dataset handling. Figure 1 demonstrates the concept of RTI acquisition setup. Each image is captured with a different illumination direction where the camera and object are fixed. This collection of acquired images is called the RTI dataset. RTI datasets are also known as Multi Light Image Collection (MLIC) or Single Camera Multi Light (SCML), depending on the research community [2, 3]. The second stage is called *Modelling*. It primarily processes the dataset to model the data from discrete to continuous space [4, 5, 1, 6]. The third stage is called *Application*. It involves utilizing the modeled data to get insights about the surface [7, 8, 9, 10, 11, 12]. For example, to study the geometry of a surface, Normal maps created from modeled RTI data can be one application that can further help in the deeper investigation of surfaces [13]. Another application can be interactive visualization of surfaces in continuously varying illumination [14, 15]. RTI has profound applications in CH of which change detection and digitization are the most popular ones [16, 17].

CVCS2024: the 12th Colour and Visual Computing Symposium, September 5–6, 2024, Gjøvik, Norway
muhammad.a.khawaja@ntnu.no (M. A. Khawaja); sony.george@ntnu.no (S. George); Franck.Marzani@u-bourgogne.fr (F. Marzani); jon.hardeberg@ntnu.no (J. Y. Hardeberg); alamin.mansouri@u-bourgogne.fr (A. Mansouri)
https://www.ntnu.edu/employees/muhammad.a.khawaja (M. A. Khawaja); https://www.ntnu.edu/employees/sony.george (S. George); https://imvia.u-bourgogne.fr/index.php/membres/marzani-franck/ (F. Marzani); https://www.ntnu.edu/employees/jon.hardeberg (J. Y. Hardeberg); https://imvia.u-bourgogne.fr/index.php/membres/mansouri-alamin/ (A. Mansouri)
0009-0003-6820-2792 (M. A. Khawaja); 0000-0001-8436-3164 (S. George); 0000-0003-0963-1565 (F. Marzani); 000-0003-1150-2498 (J. Y. Hardeberg); 0000-0001-9054-3719 (A. Mansouri)

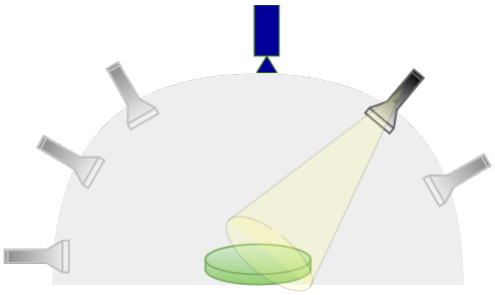


Figure 1: RTI setup: The camera is fixed at the top and the object is kept stationary at the bottom. The images are captured in series with varying illumination directions

The main idea behind RTI is exploiting the reflectance information of the surface under varying illumination. Illumination is incident light of some wavelength (visible spectrum) that hits the surface. Part of the light is absorbed by the surface and part of the light is reflected. If most of the reflected light is concentrated in one direction, it is called specular scattering. Conversely, if the light is more uniformly reflected in all directions it is called a diffuse or lambertian scattering. The comprehensive dataset of reflectance measurements of a surface point, obtained from different light directions as observed by a fixed optical sensor (camera) in a Reflectance Transformation Imaging (RTI) setup, is termed as the reflectance profile. Figure 2 demonstrates this reflectance model. Reflectance modeling is a well-studied problem in computer graphics and notable models are Phong Reflectance Model [18], Cook-Torrance Reflectance Model [19] and modern deep-learning based reflectance models like NeRF [20].

The reflectance profile captures how the appearance of a surface point changes as the light direction varies, while the viewing angle remains constant. This concept is closely related to the Bidirectional Reflectance Distribution Function (BRDF) used in computer graphics. BRDF explains how light is reflected for all possible combinations of incoming light direction, viewing direction and wavelength whereas RTI is focused on a fixed viewing direction and wavelength with only variable incoming light direction. BRDF therefore has many applications in color analysis in computer graphics [21] and surface classification in digital imaging [22] however BRDF measurement is a complex task and is often very difficult to do for many CH objects.

Classification in computer vision has been revolutionized in this decade due to recent advancements in deep learning methods [23, 24, 25]. Most of the classification is based on the geometrical features of the object. The question we ask ourselves is, while the existing methods are very good at identifying objects, can these methods provide information about the composition of the object, its tactile properties, and other physical characteristics? Humans can predict how an object feels to the touch and comprehend its properties from an image. For example, we can identify the material of the table; we can predict the feeling of touch, anticipate its tactile sensation, and understand the state of the table if it's dirty or clean from an image. Our brains have the remarkable capacity to infer the physical attributes of surfaces from visual input, and humans develop this skill from a young age.

RTI can provide an extra dimension of reflectance in the data, making it unique relative to other imaging modalities. This dimension can capture the reflectance of the surface and can also be used to estimate BRDF parameters. This BRDF can be estimated by various surface reflectance models such as the Ward model [26] shown in figure 2 which models diffuse as well as specular surfaces and can provide surface understanding beyond identification. This can help us gain insights into the material properties and conditions of the surface under study. The classification of surfaces with respect to their reflectance properties can also help in optimizing the acquisition process in RTI. RTI acquisition is conventionally done manually using homogenous, equally spaced light directions, pre-planned without using any surface information. This can produce redundant light directions, potentially leading to huge-size RTI datasets. The optimization of light directions can save a lot of resources and time for acquiring RTI datasets [17, 27, 28].

We propose to exploit the illumination variation property of RTI to classify different areas in the

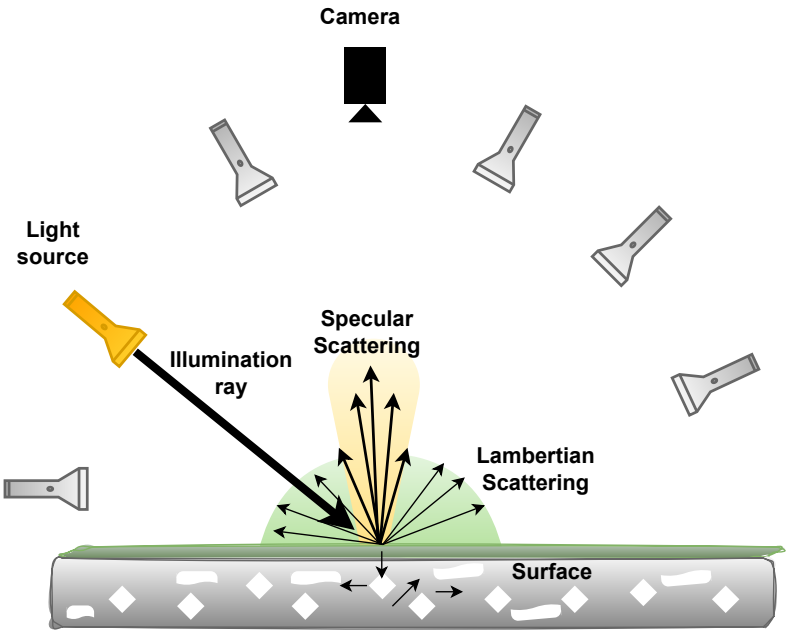


Figure 2: The illustration depicts the interaction of illumination direction with a surface, showing both lambertian or diffuse scattering and specular scattering from a light source as shown in the RTI setup. Some light is also absorbed by the surface. This interaction varies with the change in incident illumination direction

surfaces. The differences in reflectance properties within a surface could be caused by various factors. For example, rust can alter the reflectance profiles of surfaces, and dust or dirt can also affect the reflectance profile of a surface. Developing a tool to classify surfaces using RTI can aid in the study, detection, and investigation of changes or degradation. Surface classification can also lead to material identification. Material information is extensively exploited by curators, restorers, conservators, and researchers to gain insights into cultural heritage objects. Our framework's initial outcome is identifying interesting features and anomalies from the RTI data by leveraging self-supervised learning techniques like K-means and Self-Organizing Maps (SOMs). This framework can be later integrated to RTI visualization tools for interactive visualization and aims to provide a powerful tool for the analysis and visualization of surface characteristics. This tool will be valuable for applications in cultural heritage preservation, industrial quality control, and beyond, enabling users to gain comprehensive insights into the surfaces they study.

Section 2 familiarizes with the related work and section 3 presents the methodology designed for classification. The paper then introduces the dataset in section 4 used in this study, followed by section 5, which discusses the experiments conducted using the proposed methodology. Finally, section 6 presents the conclusions drawn from this study. The code for this work is available at github[1].

2. Related Work

Classification of surfaces is a complex task. Multiple modalities have been used to learn and understand the surfaces. Some related works in this field use texture with visual modality to classify surfaces using deep learning [29]. Surfaces tend to have a unique texture, which can serve as critical information in classification. A computational imaging method was presented for raw material classification using features of Bidirectional Texture Functions (BTF) in [30]. They proposed to learn discriminative illumination patterns and texture filters to directly measure optimal projections of BTFs for the classification of surfaces. The multispectral polarimetric imaging technique captures both the spectral and polarimetric

[1]https://github.com/akhawaja2014/CVCS2024_ClassificationOfSurfaces

properties of the light, adding dimensions to the spatial intensity that is normally acquired. This provides unique and discriminatory information that can help in material classification. This classification method based on multispectral polarimetric BRDF characteristics is proposed in [22]. In particular, near-infrared (NIR) range wavelengths were used for materials classification [31]. A hyperspectral imaging system was developed in [31] to classify and sort papers according to their material quality. Another paper proposed a Deep Convolutional Neural Network (DCNN)-based 60-GHz radar material classification system that uses images from the radar sensor as input features [32]. A classification accuracy of above 97 % was achieved while classifying 1875 radar images from ten different materials. These materials included books, chocolate cream, glass, the aluminum surface of a laptop, wood, wool, carpet, tile, laminate, and water. However, all these techniques are expensive because they utilize specific hardware like radar, multispectral cameras, hyperspectral cameras, or texture measurement tools and are not trivial methodologies. Furthermore, studying the object's state (whether the object is clean or dirty, etc.) has not yet been explored in these above-mentioned methods.

Some inspiration comes from the medical field. The classification of ECG signals using machine learning has been an important milestone in the diagnostic and medical analysis field [33, 34]. The ECG signals have hidden properties and profiles for a certain class of disease [35, 36]. Rastgoo *et al.* [37] proposed a classification framework for melanoma lesions using sparse-coded features and Random Forests. Melanoma is a type of skin cancer that can be treated if diagnosed.

Another source of inspiration comes from multispectral and hyperspectral image analysis. This type of analysis and classification has led to numerous applications in cultural heritage, medical imaging, quality control, and industrial inspection. Mandal *et al.* used unsupervised learning to classify pigments on Edvard Munch's self-portrait painting by utilizing hyperspectral data [38]. The classification of pigments using hyperspectral imaging provides valuable insight for understanding the painting as well as for its conservation [39, 40, 41, 42]. It also offers opportunities to estimate the era of paintings and identify forgery in paintings. Devassy *et al* proposed unsupervised clustering to classify hyperspectral data of paper, which has significant applications in document forgery investigations [43]. Another important technique for pigment identification in CH is Near Infrared Imaging (NIR). NIR imaging reveals information about the chemical composition and physical properties of pigments in CH which can also be used for surface analysis and classification [44, 45].

3. Methodology

In this paper, we propose a comprehensive methodology to classify the surfaces with respect to reflectance profiles using RTI. We divide our methodology into four major parts. The first part is data acquisition and processing. The second part includes modeling data using machine learning algorithms. The third part consists of result extraction and mapping into the spatial domain for interpretation. The final part is a visualization of results using a segmentation mask on the reference image. Figure 3 demonstrates the pipeline of the methodology.

This study primarily focuses on qualitative data (such as anomalies, degradation, rust, etc). Our objective is to identify the surfaces that are otherwise not identifiable as distinguished ones based on their appearance. This is a latent property and reasons for the change in reflectance profile might include anomaly, rust, material degradation etc. It is also important to note that this methodology is for Visible Spectrum (ViS) imaging. We have not utilized any multispectral sensor.

The virtual and real RTI both have been utilized to test our methodology. We started with a virtual RTI dataset simulated in Blender [46]. The Blender RTI toolbox mimics the RTI setup where a stationary object is illuminated from a desired number of illumination directions. This virtual approach allows us to precisely control the illumination conditions and surface properties, providing a clear and noise-free dataset for initial algorithm development and testing. The use of Blender enables us to simulate various material properties and reflectance behaviors, which are crucial for validating our classification framework in a controlled environment before applying it to real-world data. The real and simulated datasets were used to train the Self-Organizing Maps (SOMs) and K-means from scratch, ensuring that

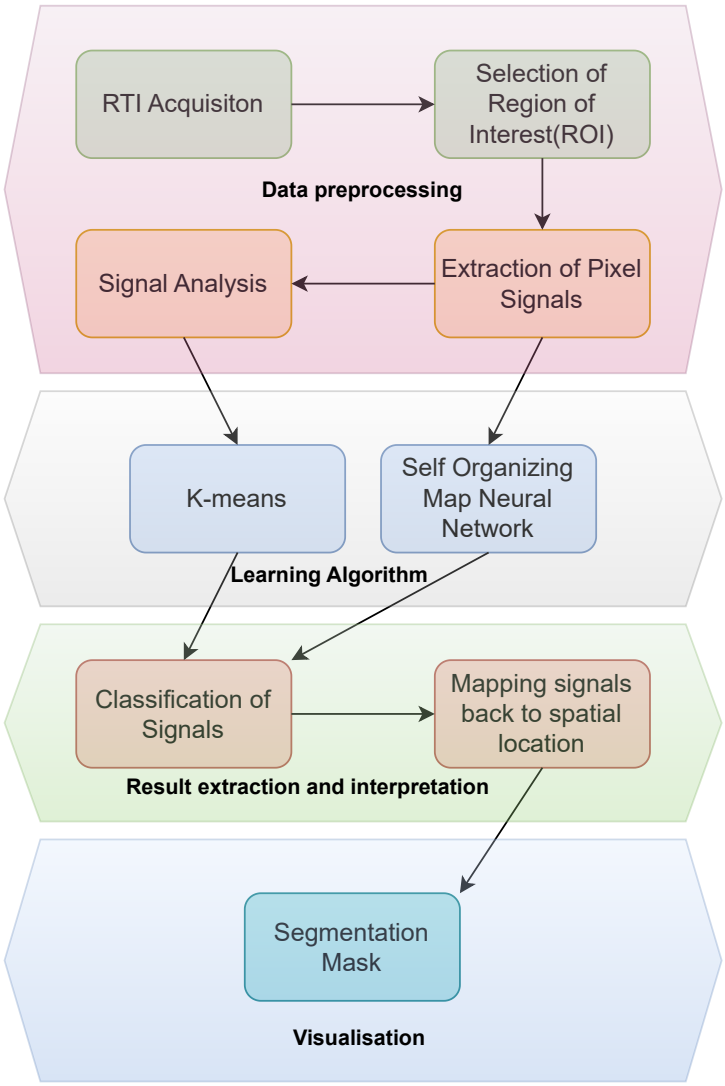


Figure 3: This flowchart outlines the step-by-step process of signal classification. It begins with RTI Acquisition and the selection of the Region of Interest (ROI), followed by the extraction of pixel signals. The signals then undergo analysis using methods such as K-means and SOM. Post-classification, the signals are mapped back to their spatial locations, resulting in a segmentation mask for visualization.

the models could correctly classify surface reflectance profiles. No pre-trained SOMs were used in this study. Section 4 explains more about the dataset. This is real-time training for each dataset and is self-supervised without the need for any labeled data.

The first step in our methodology is signal extraction. In this step, we acquire the Multi Light Image Collection (MLIC) data from a simulated environment or real acquisition in the RTI dome that is available in the lab. The acquired data needs to be pre-processed. We reduce the sample space by extracting the Region of Interest (ROI) from the dataset using a reference image. Figure 4 demonstrates the signal extraction process.

3.1. K-Means

The second and most important part of the methodology is the machine learning methods for the classification of surfaces. Initially, we used K-means clustering to classify the signals based on their standard deviation. K-means is a machine learning algorithm that divides the n-dimensional data into k-clusters using variance inside the dataset [47, 48]. While it works well for many cases, it can be vulnerable to wrongly classifying the surfaces because two different signals might have the same

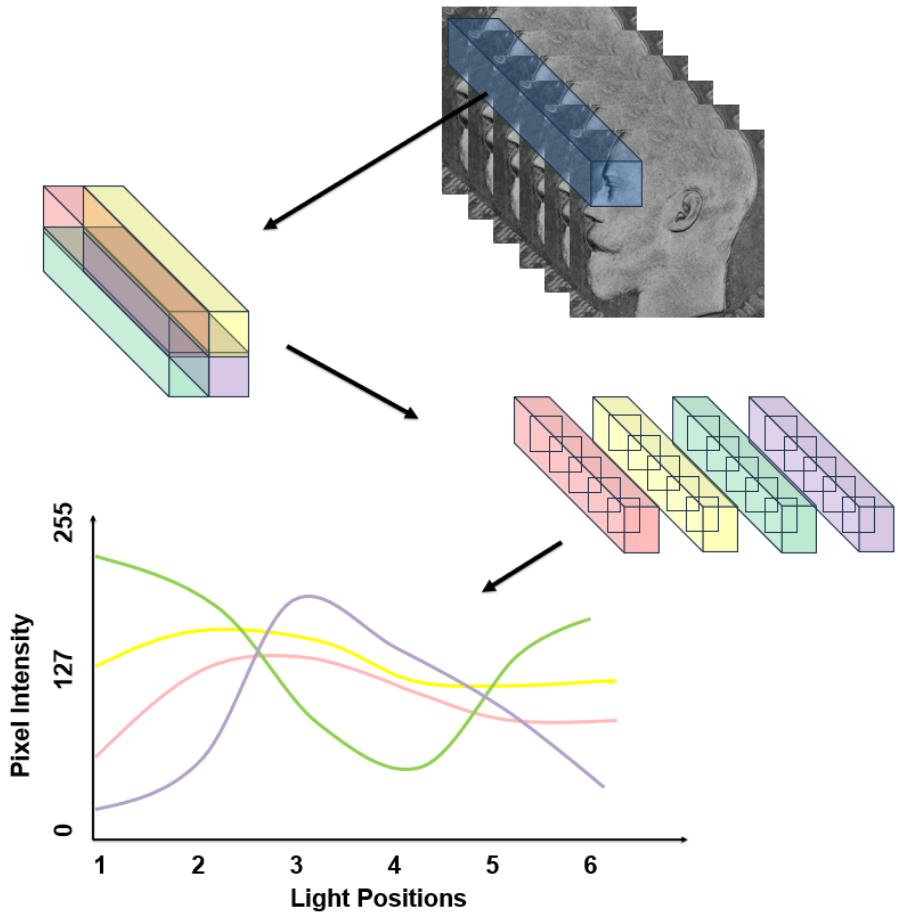


Figure 4: Demonstration of signal extraction method. The first step is to select the region of interest. Each pixel from the same location is extracted across the MLIC. The intensity value that changes with varying illumination is plotted against the light positions. The pixel signal consists of a list of varying pixel values across the MLIC for the same spatial location.

standard deviation. This motivated us to utilize SOM neural networks. The K-means algorithm that we employed for our classification is explained in pseudocode 1.

3.2. Self Organizing Maps (SOMs)

SOM's are inspired by biological neural systems. They are an abstract mathematical model of topographic mapping from visual sensors to cerebral cortex[49]. SOMs are a type of artificial neural network that is trained using unsupervised learning to produce a low-dimensional discretized representation of the training samples from input space [50, 51]. Figure 5 demonstrates the vanilla architecture of SOMs.

SOM's are trained using competitive learning rather than the error-correction learning (e.g., backpropagation using gradient descent) conventionally used by other artificial neural networks like ResNets[24], VGG [52], and Tranformers [53, 54] etc. SOM's work in two phases. The first one is training and the second one is mapping. Initially, the training phase learns from the dataset or input space to create a simplified, lower-dimensional version of *"map space"*. Finally, this established map space is used to classify new input data in the mapping phase. The SOM algorithm is summarized in pseudocode 2 adapted from [55, 49]

The third part of the algorithm is the extraction of results. The signals that are classified with their surface labels are mapped back to the spatial space of the image. Finally, a segmentation mask is built based on labels and visualized in the fourth and last part. The section 5 explains and demonstrates the experiments and results by applying the discussed methodology.

Algorithm 1 Classifying RTI signals using K-means with Standard Deviation

Input: Signal Standard deviation dataset $X = [x_1, x_2, \ldots, x_n]$, number of clusters k, threshold ϵ, iterations η
Output: Classification of signals into cluster and centroids
Initialize cluster centroids $\mu_1, \mu_2, \ldots, \mu_k \in \mathbb{R}^n$ randomly
repeat
    for each signal $x^{(i)}$ **do**
        $C_j = \left\{ x_i : \|x_i - \mu_j\|^2 \leq \|x_i - \mu_l\|^2 \text{ for all } l = 1, 2, \ldots, k \right\}$ ▷ C_j is set of data points assigned to μ_j
    for each cluster j **do**
        $\mu_j = \frac{1}{|C_j|} \sum_{x_i \in C_j} x_i$ ▷ Update cluster centroids
    Convergence Check:
    if $||\mu_j^{new} - \mu_j^{old}|| < \epsilon$ **then**
        break
        if $iter == \eta$ **then**
            break
until convergence
return cluster assignments and centroids

4. Dataset

We have prepared some in-house data, and other data has been obtained from published works [56, 57]. For the acquisition of MLIC, we selected the ring setup (varying azimuth, fixed elevation) due to its capability to showcase unpredictable patterns in pixel intensity because of surface characteristics. This lets us focus more on the data depending on surface characteristics. The data is explained in the following sub-sections.

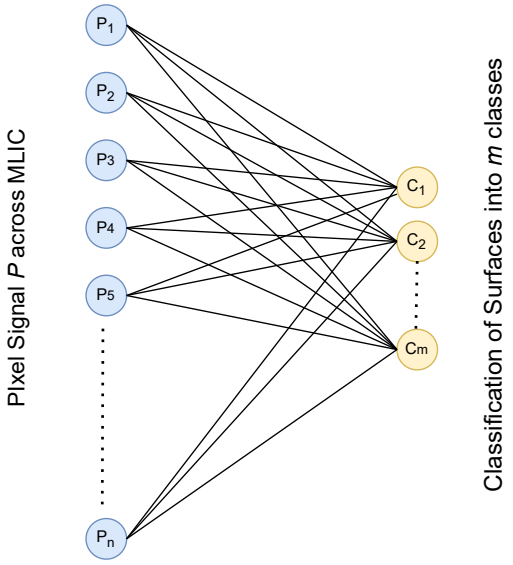


Figure 5: SOMs map high-dimensional data into a lower-dimensional space while preserving the topological structure of the data. The n-pixel signals are classified into m-classes

Algorithm 2 SOM for RTI signal classification

Input: RTI Signals $\mathbf{X} = [x_1, x_2, \ldots, x_n]$, iteration t, learning rate α
Output: Classification of signals into m classes
Initialize the weight vectors $\mathbf{W} = [w_1, w_2, \ldots, w_m]$ randomly
for $t = 1$ to T **do**
    At each iteration t, present an input signal $x(t)$, and select the winner using competition.
    $\nu(t) = \arg\min_{k \in \Omega} \|\mathbf{x}(t) - \mathbf{w}_k(t)\|$
    Update weights of the winner and its neighbors.
    $\Delta \mathbf{w}_k(t) = \alpha(t)\eta(\nu, k, t)\,[\mathbf{x}(t) - \mathbf{w}_\nu(t)]$ ▷ Ω is a set of neuron indexes
    where:
    $\eta(\nu, k, t) = \exp\left[-\frac{\|\mathbf{r}_\nu - \mathbf{r}_k\|^2}{2\sigma(t)^2}\right]$ ▷ neighbourhood function
    Learning rate decreases monotonically and satisfies the following equations:
    $0 < \alpha(t) < 1$
    $\lim_{t\to\infty} \sum \alpha(t) \to \infty$
    $\lim_{t\to\infty} \alpha(t) \to 0$
return updated $W = [w_1, w_2, \ldots, w_m]$
$Y = W(X)$
$Classes = vec2ind(Y)$

4.0.1. Coin of Emperor Nicolas

The first surface consists of a cultural heritage object. It is a coin called the *"Coin of Emperor Nicolas II"* from the year *1897 AD*, depicted in figure 6. This coin's 3D model is sourced from Sketchfab under a Creative Commons license and was used in our study [56]. We used Blender software to capture the RTI images.

Figure 6: Coin of Emperor Nicolas II. It is 1 Rouble coin that was in circulation from the year 1895 to 1915 in what is now modern-day Russia. This coin weighs $20g$, has a diameter of $33.65mm$, a thickness of $2.7mm$, and is made of silver. Source: [58]

4.0.2. Rust Course Metal

The second surface we used in our experiments was *"Rust Course Metal"*, which is obtained from a dataset published by [57]. It consists of a metallic surface with some rust on it. The figure 7 shows the metal. This surface is important to understand how the reflectance properties of metals change when rust comes on them. This has significance in industrial quality control and material inspection.

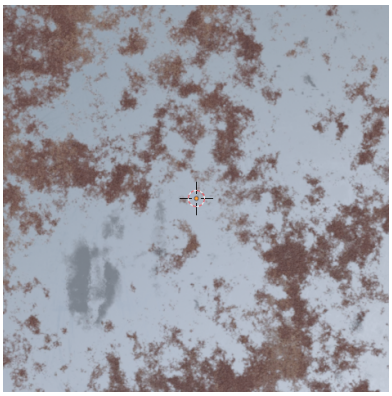
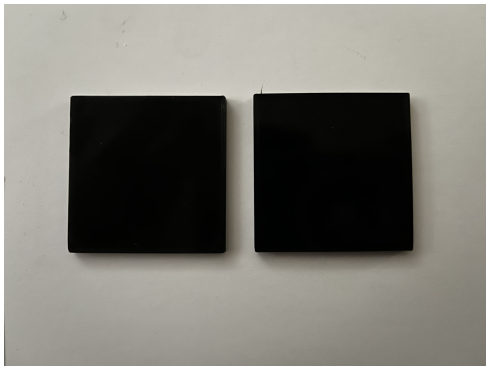


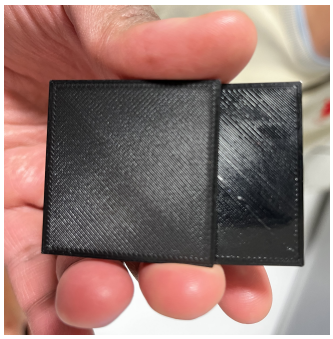
Figure 7: Rust Course Metal. Source: [57]

4.0.3. Two Black Surfaces

This dataset was prepared to compare the algorithm's performance in simulated data and real data. Two black surfaces with the same color perception to the human eye but different material properties were prepared by 3D printing. We have used the grayscale camera, primarily a one-channel camera, to acquire its MLIC dataset. Figure 8 demonstrates the real *"Two black Surfaces"*.



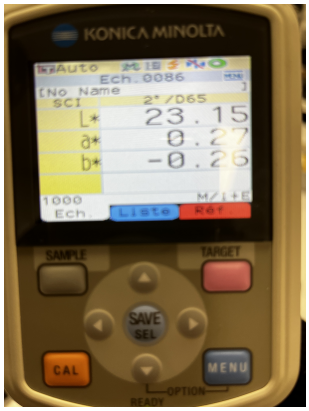
(a) The surface on the left side is PLC and the one on right side is Nylon.



(b) Demonstration of varying reflectance profiles of surfaces

(c) Color measurement values of PLC.

(d) Measuring L*a*b* values of both surfaces.



(e) Color measurement values of Nylon.

Figure 8: *Two black surfaces (real)*: It can be seen that these two surfaces have similar color. However, they are made of different materials and hence possess different reflectance properties. Nylon is shinier and PLC is more matt. The texture on them is a fabrication constraint.

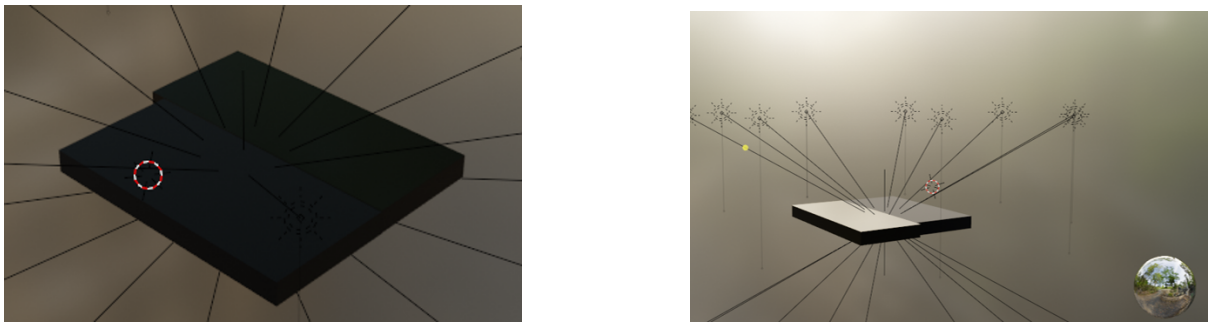


Figure 9: *Two black surfaces (simulated):* It can be seen that while these two surfaces have same color but they are made of different materials and hence possess different reflectance properties. The images are simulated in blender software.

The synthetic/simulated "Two black surfaces" were also created in blender software by putting the same constraint of the same color and different material. The metallic and roughness properties were kept different for both surfaces. Figure 9 shows the two black surfaces on the blender. Since it is a simulated environment, there was no fabrication constraint and hence the surface did not have any texture lines like the real "Two black surfaces".

5. Experiments and Results

This section employs the methodology discussed in the section 3 to experiment on data explained in section 4. The results are obtained and discussed.

The first experiment was done on the *"Coin of Emperor Nicolas"*. Since the coin is old and a cultural heritage object, it has undergone various aging processes, and its surface is no longer homogenous. We aim to differentiate the surface within the coin based on its illumination variation information. This can further help cultural heritage experts understand the level of rust or other degradation process that can be present and not apparent otherwise. We used 27 distinct illumination directions at a constant elevation of 30° and variable equally spaced azimuth to generate 27 simulated images in the blender. Figure 10 demonstrates and explains the results. The same experiment was repeated with the K-means algorithm, and results are shown in figure 11.

The second experiment was performed on the *"Rust Course Metal"*. We used 10 distinct illumination directions with a constant elevation of 30° and equally spaced variable azimuth angles to generate 10 simulated RTI images. Figure 12 explains the results. Since it is a synthetic flat surface with symmetric reflectance profile, ring acquisition was not helpful in getting illumination variation data. We found that using a varying elevation acquisition, where lights are positioned at different elevation angles, would be more suitable for obtaining meaningful reflectance data from this type of surface.

The third experiment was performed on the real *"Two Black Surfaces"*. The two black surfaces were assembled in the RTI dome to get MLIC. We captured 24 distinct images from 24 illumination directions with a constant elevation of 45° and variable equally spaced azimuth. The camera in the RTI dome is very high quality and produces 4000 by 6000-pixel images. Hence, it was difficult to process big regions of interest. While the algorithm struggles to classify real data, as shown in figure 13, it was able to differentiate one surface from another to some extent.

The fourth experiment was performed on the synthetic *"Two Black Surfaces"*. We used 10 distinct illumination directions with a constant elevation of 30° and equally spaced variable azimuth angles to generate 10 simulated RTI images. This surface is discussed in section 4 and figure 9. The algorithm classifies the two different materials in *"Two Black Surfaces"* as shown in figure 14 because it is a clean synthetic surface with a distinguished reflectance profile.

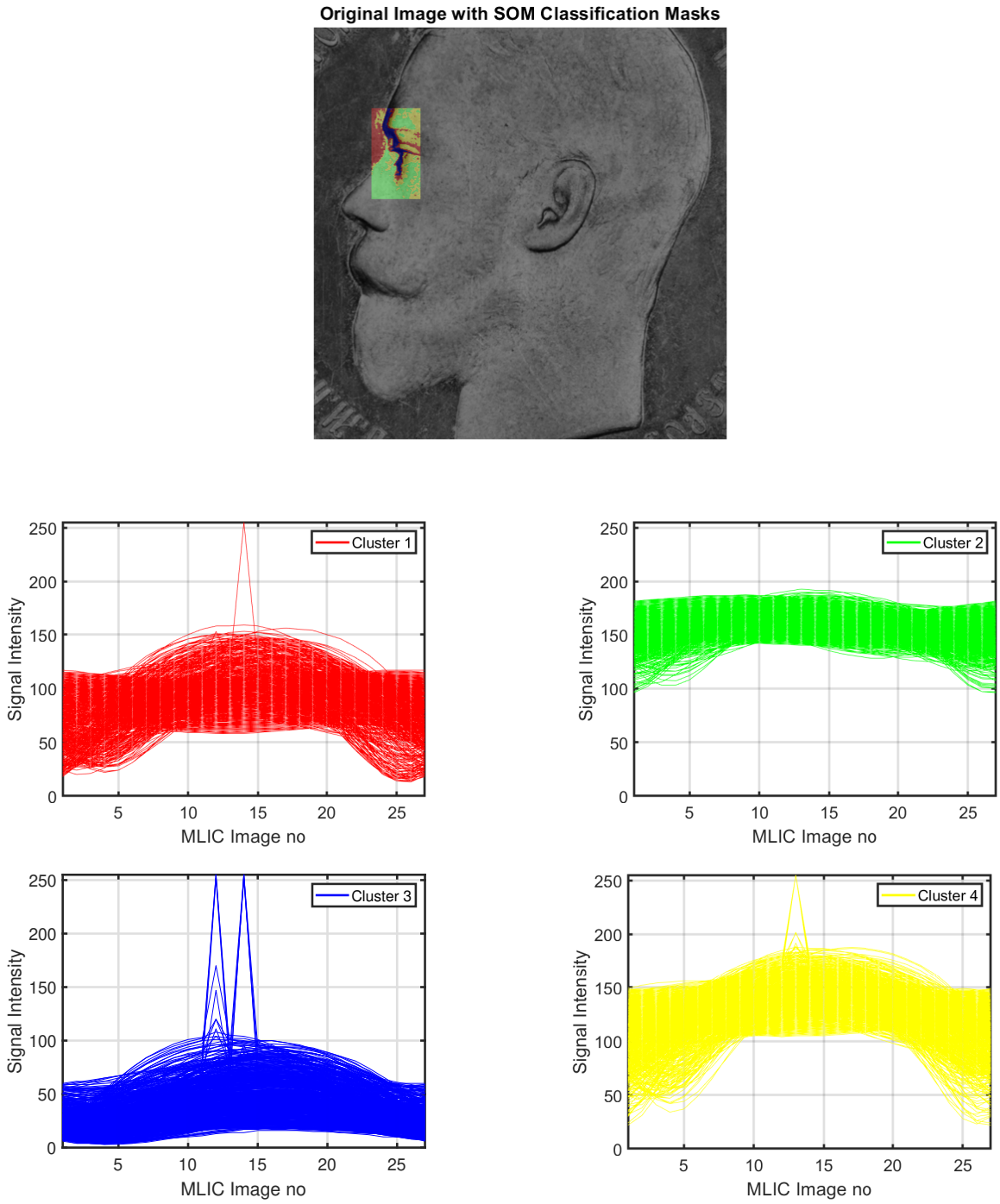


Figure 10: It can be seen that the pixels have been classified according to their reflectances w.r.t varying illuminations. SOMs learn the pattern in the signals and classify them accordingly. The yellow segmented area represents a bright surface type whereas green and blue are relatively specular surface types considering their sensitivity to light directions. The red surface type is a relatively dark background. This can help cultural heritage experts identify potential degradation that might not be apparent on the surface. Non-experts might struggle due to the complexity of the variations in reflectance.

6. Conclusion

This paper presents an approach to surface classification using RTI data. The suggested method uses SOM and K-means, two self-supervised learning algorithms, to classify surfaces according to their reflectance profiles. Surface classification using RTI has great potential for several applications, such as industrial quality control and the preservation of cultural heritage. We believe surface classification is

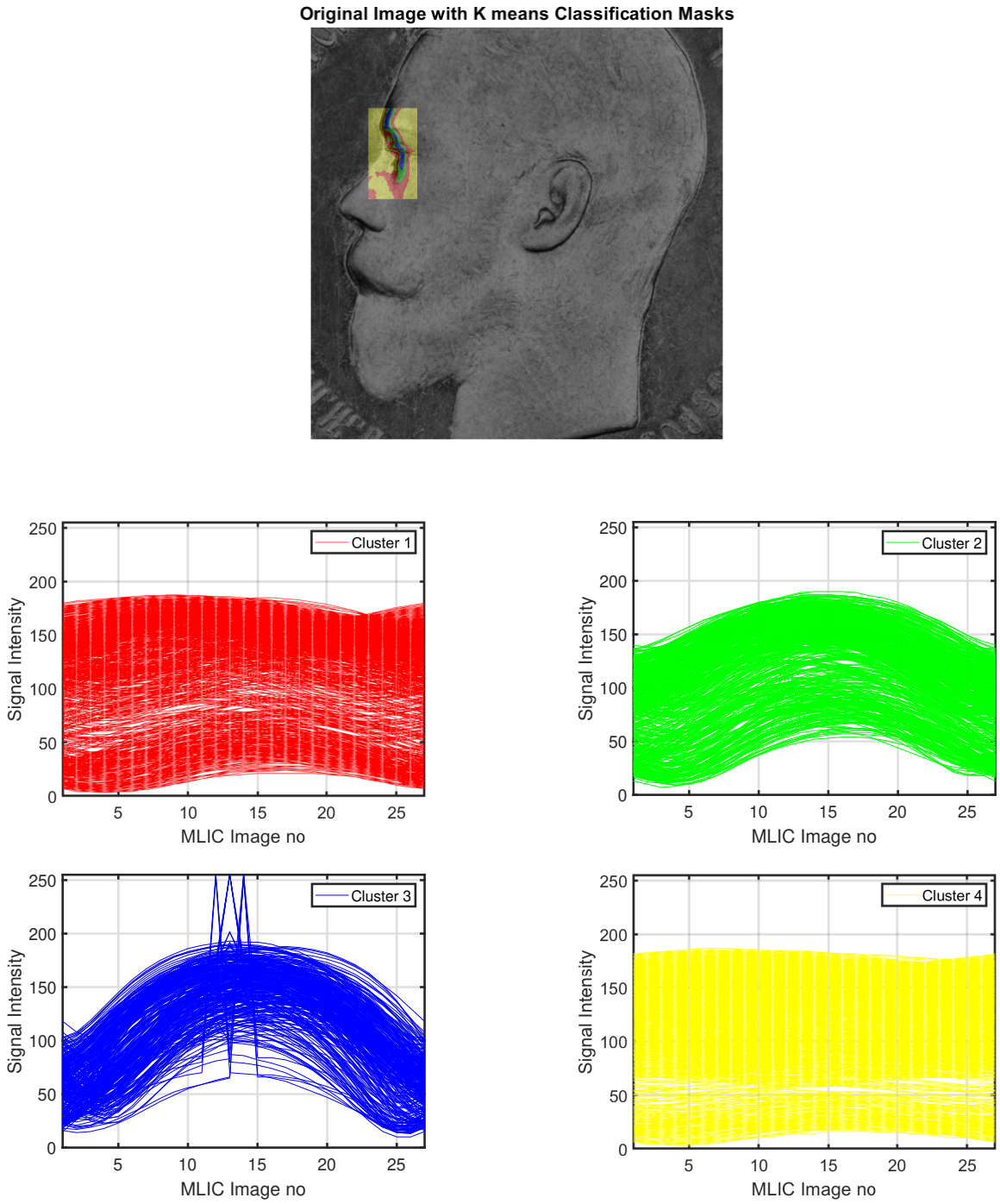


Figure 11: K-means classified the surface using standard deviation in the signal. The pixels with similar reflectance profiles are clustered together. This segmentation mask tells what part of the surfaces have illumination variation (for example, Blue and Green classes here). However, it can be seen that means is misclassifying dark background (yellow class) which is dark with a bright part of the face since they have similar standard deviation.

the first step to material classification using RTI.

Machine learning plays a crucial role in surface classification problems. We used K-means classification initially, but it had drawbacks because signals can be mistakenly misclassified since K-means classifies signals based on standard deviation. It does not learn the patterns of the signal. This resulted in the use of artificial neural networks known as SOM neural networks, which is a self-supervised learning method and offers a better low-dimensional representation of data. The examination of coin of Emperor Nicolas II provides us with valuable information about the reflectance properties of the

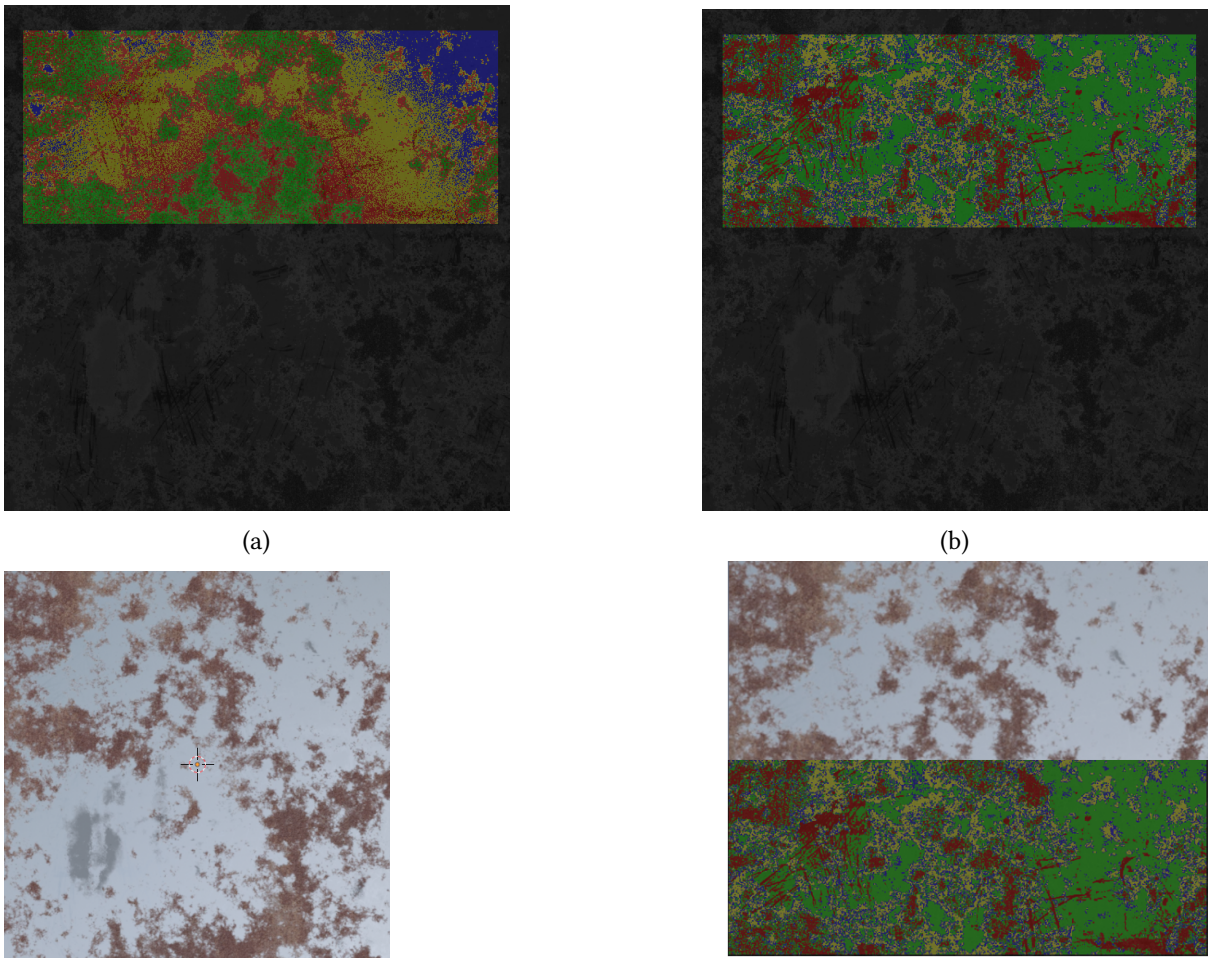


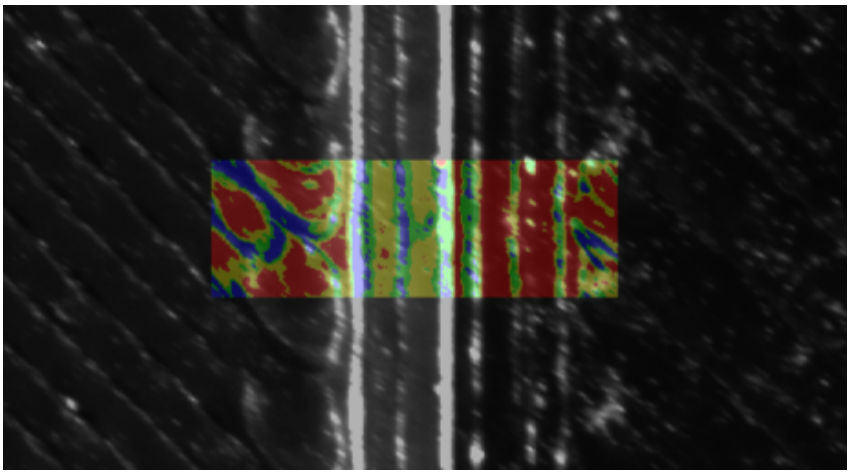
Figure 12: The sub-figure (a) shows the classification of the surface using K-means, sub-figure (b) shows classification using SOM neural network, and sub-figure (c) shows the original color image of the surface. It can be observed that SOMs perform better than K means. SOMs are able to detect scratches in the surface which appear like lines in sub-figure (b). Sub-figure (d) shows the compares the same segmented part in subfigure (c) with its SOM segmented part.

surfaces. This methodology is designed for 2D plain surfaces and struggled with real *"Two Black Surfaces"* whose geometry was not completely flat due to 3D printer fabrication constraints. The algorithm was identifying shadow of elevations as matt surface which was the wrong classification as discussed in section 5. However, the algorithm was still able to classify surfaces according to their reflectance profiles. We observed that the K-means algorithm does not perform as well as the SOM neural networks in our experiments, however this observation is primarily based on visual inspection of the classification results. This qualitative assessment, though not quantified by classification mismatch metrics, can be supported by expert validation of the results.

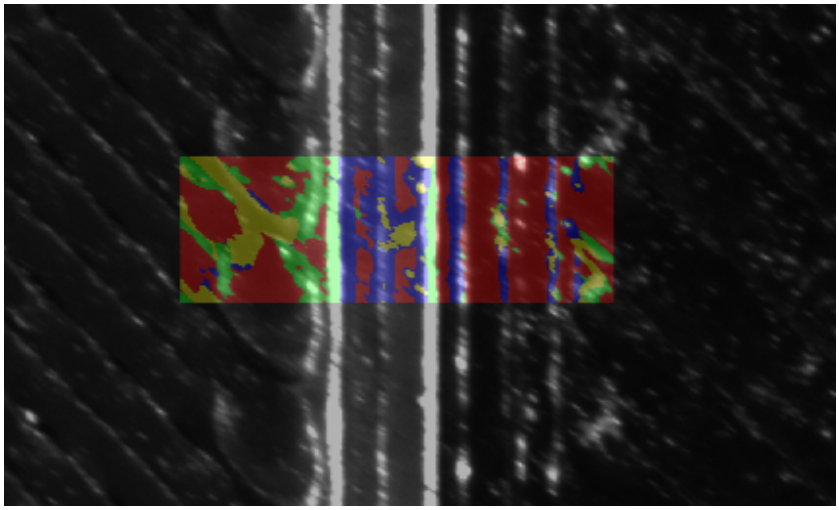
Our approach's motivation is to exploit the illumination variation data to better understand the surfaces and discover their anomalies. This is the prime reason we are using the self-supervised method, and since we are exploring the surface, we do not have the ground truth to compare our algorithm.

In summary, the experiments presented here contribute to the advancement of techniques for surface analysis, with implications for fields such as cultural heritage preservation and material characterization.

In the future, our research will focus on refining and creating methodologies to address computational challenges, expanding the scope of materials analyzed and addressing geometrical 3D properties of the surfaces in the algorithm. We plan to extend this classification model to use RGB images. RGB images offer extra color information that can help the model better distinguish between different materials,



(a) Segmentation of two black surfaces using K-means. It can be seen that the K-means algorithm has classified the right side (PLC matt surface with red) mostly correctly but is struggling with classifying the relatively specular left side (Nylon). This is due to the involvement of texture geometry.



(b) Segmetation of two black surfaces using SOM neural network. It can be seen that the SOM algorithm mostly classifies the right PLC side as one class (red) based on its matt response. However, it has misclassified some parts of the left side surface because it mistakes shadow a matt surface due to texture geometry.

Figure 13: The gray lines between them signify the separation. The left side surface is more shiny while the right side surface is more matt. The fabrication constraints of texture lines have complicated the reflectance responses, and this makes the algorithm vulnerable to wrongly classifying the surface.

potentially improving accuracy. We also plan to extend this framework for 3D RTI acquisition. This means the algorithm will be extended for RTI acquisitions with variable elevations and variable azimuth angles. This can provide invaluable reflectance variation information.

Acknowledgments

We would like to express our deepest gratitude to ISITE-BFC project. (Initiatives Science Innovation Territoire Economie en Bourgogne-Franche-Comté) of Université de Bourgogne for their unwavering support and funding throughout the course of this research.

Figure 14: The synthetic *"Two Black Surfaces"* are classified correctly using SOM. The red class shows the first surface and the green class shows the second surface.

References

[1] T. Malzbender, D. Gelb, H. Wolters, Polynomial texture maps, in: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '01, Association for Computing Machinery, New York, NY, USA, 2001, p. 519–528. URL: https://doi.org/10.1145/383259.383320. doi:10.1145/383259.383320.

[2] R. J. Woodham, Photometric method for determining surface orientation from multiple images, Optical Engineering 19 (1980) 191139. URL: https://doi.org/10.1117/12.7972479. doi:10.1117/12.7972479.

[3] V. Vanweddingen, H. Hameeuw, B. Vandermeulen, C. Vastenhoud, L. Watteeuw, F. Lemmers, A. V. der Perre, P. Konijn, L. V. Gool, M. Proesmans, Pixel+: integrating and standardizing of various interactive single-camera, multi-light imagery, in: P. Schelkens, T. Kozacki (Eds.), Optics, Photonics and Digital Technologies for Imaging Applications VI, volume 11353, International Society for Optics and Photonics, SPIE, 2020, p. 113530G. URL: https://doi.org/10.1117/12.2555685. doi:10.1117/12.2555685.

[4] P. Gautron, J. Krivánek, S. N. Pattanaik, K. Bouatouch, A novel hemispherical basis for accurate and efficient rendering., Rendering Techniques 2004 (2004) 321–330.

[5] G. Pitard, G. Le Goïc, A. Mansouri, H. Favrelière, S.-F. Desage, S. Samper, M. Pillet, Discrete modal decomposition: a new approach for the reflectance modeling and rendering of real surfaces, Machine Vision and Applications 28 (2017) 607–621.

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

[7] A. Zendagui, G. Le Goïc, H. Chatoux, J.-B. Thomas, Y. Castro, M. Nurit, A. Mansouri, Quality assessment of dynamic virtual relighting from RTI data: application to the inspection of engineering surfaces, in: Fifteenth International Conference on Quality Control by Artificial Vision, volume 11794, SPIE, 2021, pp. 94–102.

[8] H. Mytum, J. R. Peterson, The application of reflectance transformation imaging (RTI) in historical archaeology, Historical Archaeology 52 (2018) 489–503.

[9] A. Zendagui, G. Le Goïc, H. Chatoux, J.-B. Thomas, P. Jochum, S. Maniglier, A. Mansouri, Reflectance transformation imaging as a tool for computer-aided visual inspection, Applied Sciences 12 (2022) 6610.

[10] G. Pitard, G. Le Goïc, A. Mansouri, H. Favrelière, M. Pillet, S. George, J. Y. Hardeberg, Robust anomaly detection using reflectance transformation imaging for surface quality inspection, in: Image Analysis: 20th Scandinavian Conference, SCIA 2017, Tromsø, Norway, June 12–14, 2017, Proceedings, Part I 20, Springer, 2017, pp. 550–561.

[11] G. Earl, P. Basford, A. Bischoff, A. Bowman, C. Crowther, J. Dahl, M. Hodgson, L. Isaksen, E. Kotoula, K. Martinez, et al., Reflectance transformation imaging systems for ancient documentary artefacts, in: Electronic Visualisation and the Arts (EVA 2011), BCS Learning & Development, 2011.

[12] D. Sharma, M. Nurit, U. Rothenhäusler, K. Schmidt-Ott, E. Joseph, S. George, T. Lombardo, Application of reflectance transformation imaging for visualizing early signs of corrosion in historical glass corrosion, in: Archiving Conference, volume 20, Society for Imaging Science and Technology, 2023, pp. 143–148.

[13] D. A. Lewis, H. Chatoux, A. Mansouri, SFF-RTI: An active multi-light approach to shape from focus, The Visual Computer 40 (2024) 2067–2079.

[14] L. Righetto, A. Traviglia, M. De Bernardin, E. Gobbetti, F. Ponchio, A. Giachetti, Ancient coins' surface inspection with web-based neural RTI visualization, in: Optics for Arts, Architecture, and Archaeology (O3A) IX, volume 12620, SPIE, 2023, p. 12620:0D. URL: http://vic.crs4.it/vic/cgi-bin/bib-page.cgi?id='Righetto:2023:ACS'. doi:10.1117/12.2674888.

[15] L. Righetto, F. Bettio, F. Ponchio, A. Giachetti, E. Gobbetti, et al., Effective interactive visualization of neural relightable images in a web-based multi-layered framework, in: Eurographics Workshop on Graphics and Cultural Heritage, 2023, pp. 57–66.

[16] A. Siatou, M. Nurit, Y. Castro, G. Le Goïc, L. Brambilla, C. Degrigny, A. Mansouri, New methodological approaches in reflectance transformation imaging applications for conservation documentation of cultural heritage metal objects, Journal of Cultural Heritage 58 (2022) 274–283.

[17] M. A. Khawaja, S. George, F. Marzani, J. Y. Hardeberg, A. Mansouri, An interactive method for adaptive acquisition in reflectance transformation imaging for cultural heritage, in: IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), 2023, pp. 1690–1698. doi:10.1109/ICCVW60793.2023.00185.

[18] B. T. Phong, Illumination for computer generated pictures, Commun. ACM 18 (1975) 311–317. URL: https://doi.org/10.1145/360825.360839. doi:10.1145/360825.360839.

[19] R. L. Cook, K. E. Torrance, A reflectance model for computer graphics, ACM Siggraph Computer Graphics 15 (1981) 307–316.

[20] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, R. Ng, Nerf: Representing scenes as neural radiance fields for view synthesis, Communications of the ACM 65 (2021) 99–106.

[21] D. Guarnera, G. C. Guarnera, A. Ghosh, C. Denk, M. Glencross, BRDF representation and acquisition, in: Computer Graphics Forum, volume 35, Wiley Online Library, 2016, pp. 625–650.

[22] C. Chen, Y.-q. Zhao, L. Luo, D. Liu, Q. Pan, Robust materials classification based on multispectral polarimetric brdf imagery, in: International Symposium on Photoelectronic Detection and Imaging 2009: Advances in Imaging Detectors and Applications, volume 7384, SPIE, 2009, pp. 220–227.

[23] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1, NIPS'12, Curran Associates Inc., Red Hook, NY, USA, 2012, p. 1097–1105.

[24] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778. doi:10.1109/CVPR.2016.90.

[25] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, N. Houlsby, An image is worth 16x16 words: Transformers for image recognition at scale, ArXiv abs/2010.11929 (2020). URL: https://api.semanticscholar.org/CorpusID:225039882.

[26] G. J. Ward, Measuring and modeling anisotropic reflection, in: Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques, 1992, pp. 265–272.

[27] R. Luxman, M. Nurit, G. Le Goïc, F. Marzani, A. Mansouri, Next best light position: A self configuring approach for the reflectance transformation imaging acquisition process, Electronic Imaging 33 (2021) 1–7.

[28] M. A. Khawaja, S. George, F. Marzani, J. Y. Hardeberg, A. Mansouri, Can surface topography give us best light positions for reflectance transformation imaging?, Archiving Conference 20 (2023) 12–12. URL: https://library.imaging.org/archiving/articles/20/1/3. doi:10.2352/issn.2168-3204.

2023.20.1.3.
[29] H. Zheng, L. Fang, M. Ji, M. Strese, Y. Özer, E. Steinbach, Deep learning for surface material classification using haptic and visual information, IEEE Transactions on Multimedia 18 (2016) 2407–2416.
[30] C. Liu, G. Yang, J. Gu, Learning discriminative illumination and filters for raw material classification with optimal projections of bidirectional texture functions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 1430–1437.
[31] P. Tatzer, M. Wolf, T. Panner, Industrial application for inline material sorting using hyperspectral imaging in the nir range, Real-Time Imaging 11 (2005) 99–107.
[32] J. Weiß, A. Santra, Material classification using 60-ghz radar and deep convolutional neural network, in: 2019 International Radar Conference (RADAR), IEEE, 2019, pp. 1–6.
[33] S. Dalal, V. P. Vishwakarma, Classification of ECG signals using multi-cumulants based evolutionary hybrid classifier, Scientific Reports 11 (2021) 15092.
[34] J. H. Abawajy, A. V. Kelarev, M. Chowdhury, Multistage approach for clustering and classification of ECG data, Computer Methods and Programs in Biomedicine 112 (2013) 720–730.
[35] A. Gacek, Preprocessing and analysis of ECG signals–a self-organizing maps approach, Expert Systems with Applications 38 (2011) 9008–9013.
[36] M. Balouchestani, L. Sugavaneswaran, S. Krishnan, Advanced k-means clustering algorithm for large ECG data sets based on K-SVD approach, in: 2014 9th International Symposium on Communication Systems, Networks and Digital Sign (CSNDSP), 2014, pp. 177–182. doi:10.1109/CSNDSP.2014.6923820.
[37] M. Rastgoo, G. Lemaître, O. Morel, J. Massich, R. Garcia, F. Meriaudeau, F. Marzani, D. Sidibé, Classification of melanoma lesions using sparse coded features and random forests, in: Medical Imaging 2016: Computer-Aided Diagnosis, volume 9785, SPIE, 2016, pp. 73–81.
[38] D. J. Mandal, H. Deborah, S. George, J. Y. Hardeberg, Unsupervised clustering for works of art using hyperspectral imaging: A case study on Edvard Munch's self-portrait (1905), in: 2023 13th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2023, pp. 1–5. doi:10.1109/WHISPERS61460.2023.10430604.
[39] D. J. Mandal, Image Quality Assessment of Hyperspectral and Conventional Imaging for Cultural Heritage Artifacts, Phd Thesis, Norwegian University of Science and Technology (NTNU), 2815 Gjøvik, Norway, 2023.
[40] H. Deborah, S. George, J. Y. Hardeberg, Pigment mapping of the Scream (1893) based on hyperspectral imaging, in: Image and Signal Processing: 6th International Conference, ICISP 2014, Cherbourg, France, June 30–July 2, 2014. Proceedings 6, Springer, 2014, pp. 247–256.
[41] D. J. Mandal, M. Pedersen, S. George, C. Boust, Comparison of pigment classification algorithms on non-flat surfaces using hyperspectral imaging, Journal of Imaging Science and Technology 67 (2023) 1–25.
[42] H. Deborah, S. George, J. Y. Hardeberg, J. S. Ferrer, I. C. Sandu, Old man in Warnemünde (1907) colouring palette: A case study on the use of hyperspectral imaging for pigment identification, in: Color and Imaging Conference, volume 25, Society for Imaging Science and Technology, 2017, pp. 339–344.
[43] B. Melit Devassy, S. George, P. Nussbaum, Unsupervised clustering of hyperspectral paper data using t-SNE, Journal of Ismaging 6 (2020) 29.
[44] A. Dal Fovo, A. Mazzinghi, S. Omarini, E. Pampaloni, C. Ruberto, J. Striova, R. Fontana, Non-invasive mapping methods for pigments analysis of roman mural paintings, Journal of Cultural Heritage 43 (2020) 311–318.
[45] M. Attas, E. Cloutis, C. Collins, D. Goltz, C. Majzels, J. R. Mansfield, H. H. Mantsch, Near-infrared spectroscopic imaging in art conservation: investigation of drawing constituents, Journal of Cultural Heritage 4 (2003) 127–136.
[46] B. O. Community, Blender - a 3D modelling and rendering package, Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018. URL: http://www.blender.org.
[47] J. MacQueen, et al., Some methods for classification and analysis of multivariate observations, in:

Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, volume 1, Oakland, CA, USA, 1967, pp. 281–297.
[48] M. Ahmed, R. Seraj, S. M. S. Islam, The K-Means algorithm: A comprehensive survey and performance evaluation, Electronics 9 (2020) 1295.
[49] H. Yin, The self-organizing maps: background, theories, extensions and applications, in: Computational Intelligence: A compendium, Springer, 2008, pp. 715–762.
[50] T. Kohonen, The self-organizing map, Proceedings of the IEEE 78 (1990) 1464–1480. doi:10.1109/5.58325.
[51] D. Miljković, Brief review of self-organizing maps, in: 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), IEEE, 2017, pp. 1061–1066.
[52] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556 (2014).
[53] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, Advances in Neural Information Processing Systems 30 (2017).
[54] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, M. Shah, Transformers in vision: A survey, ACM computing surveys (CSUR) 54 (2022) 1–41.
[55] T. Kohonen, Self-organization and associative memory, volume 8, Springer Science & Business Media, 2012.
[56] E. Champion, H. Rahaman, Survey of 3d digital heritage repositories and platforms, Virtual Archaeology Review 11 (2020) 1–15.
[57] R. Luxman, H. Chatoux, G. Le Goïc, J. Y. Hardeberg, F. Marzani, A. Mansouri, A benchmark dataset and evaluation for best light configuration in reflectance transformation imaging, in: Archiving Conference, volume 20, Society for Imaging Science and Technology, 2023, pp. 75–81.
[58] L. Reid, S. McDougall, C. Erolin, Sketchfab: An educational asset for learning anatomy, Journal of Anatomy 236 (2020) 267.

7. Online Resources

The demo code and data is available on the *GitHub project webpage*.