International Journal of Architecture and Urbanism Vol. 08, No. 03 (2024) 479 – 496

talenta PUBLISHER

International Journal of Architecture and Urbanism

Journal homepage: https://talenta.usu.ac.id/ijau

Study of Method of Evaluating Surface's Pollution Level & Cleaning Efficiency on Decoration Building Materials

Un-Gyong O*[1], Kyong-Chon Kim[1], Jong-Chol Jo[1]

[1]*Pyongyang University of Architecture, Pyongyang, DPRK*

*Corresponding Author: shypinguo202122@yeah.net

ARTICLE INFO

Article history:
Received 24-7-2024
Revised 12-11-2024
Accepted 14-11-2024
Available online 30-11-2024

E-ISSN: 2622-1640
P-ISSN: 2622-0008

How to cite:
Un-gyong O, Kim k. C., Jo J.C. Study of Method of Evaluating Surface's Pollution Level & Cleaning Efficiency on Decoration Building Materials. International Journal of Architecture and Urbanism. 2024. 8(3):479-496.

ABSTRACT

In the paper, we analyzed synthetically the research results on methods of testing the cleaning effect on DBM(Decoration Building Materials) and suggested a new Pollution Level & Cleaning Effect Evaluation System(PCEES) on the DBM's surface which is based on image information processing by use of color quantization method and color feature extraction method with HSI (Hue Saturation Intensity) model. By applying the PCEES to many building objects, we could select the effective composition of cleaning agent and cleaning method and calculate the agent's consumption amount necessary for the building work. And we suggested the degree of cleaning effect according to the quantitative result values of evaluation on pollution level & cleaning effect to correspond with the empirical judgment of a man.

Keywords: building materials, pollution level evaluation

1. Introduction

Various types of stone, glass, and tiles with diverse colors and geometric patterns are now widely used in modern architecture due to their durability and aesthetic appeal. However, these decorative building materials often suffer surface contamination from various factors during construction and use, diminishing their decorative impact [1][2]. Consequently, significant efforts are dedicated to preventing pollution and discoloration of these materials to preserve both the quality and aesthetic of buildings, aligned with advancements in science and technology and the rising cultural standards of society. One crucial issue in this regard is the accurate assessment of pollution levels and the cleaning efficacy on building material surfaces.

Intensive research efforts on solutions to pollution and cleaning evaluation are being conducted across various countries, focusing on selecting appropriate indices for assessing cleaning effectiveness and standardizing associated testing methods [3][4]. Studies also address methodologies for applying these cleaning effect test methods across different stages, including pre-cleaning, active cleaning, and post-cleaning assessments [3][4]. Synthesizing all available cleaning effect test methods (CETM) for stone surfaces, we categorized CETM into three main approaches: instrument-based methods [3][5][6], empirical and passive methods [7], and image information processing methods. Among these, the Artificial Neural Network method within image processing has gained attention but faces limitations, as it heavily depends on color measurement instruments and extensive datasets. To address these gaps, we proposed a similarity evaluation method based on image feature retrieval and color quantization to quantitatively assess pollution levels and cleaning effects on building materials, enabling simpler field application without complex processes. Through comprehensive document analysis, we concluded that using a reduced color quantization method with an HSI model, alongside the similarity evaluation approach, offers the most effective solution for pollution level and cleaning effect estimation (PCEE). This informed the development of algorithms and a specialized computer program. We further established cleaning effect levels corresponding to quantitative PCEE results that align with human empirical evaluations. Ultimately, we developed a scientific and theoretical framework for multidimensional modeling of the entire cleaning process, encompassing pollution source identification, pollution area calculation, cleaning agent selection, and consumption estimation within our evaluation program.

In this study, the theoretical approach to the Pollution Level and Cleaning Effectiveness Evaluation System (PCEES) for decorative building materials is based on image information processing using color quantization in the HSI (Hue, Saturation, Intensity) color model. This system integrates quantitative assessments of pollution levels and cleaning effectiveness on building material surfaces with an approach that empirically aligns with human visual perception. The emphasis on selecting the HSI color model over other models, such as RGB or Lab, is supported by HSI's ability to differentiate colors accurately according to human perception, making it suitable for evaluating subtle color changes due to pollution or cleaning processes [8][9].

The theoretical foundation of this system also includes the importance of image quantization within the HSI model, dividing colors into 96 segments to simplify data without sacrificing perceptual accuracy. PCEES calculates similarity between images of pre-cleaned, polluted, and post-cleaned surfaces, yielding quantitative values that indicate color changes associated with cleanliness levels [10][11]. This system combines histogram-based similarity methods, providing results that reflect color changes objectively without requiring complex physical measurements. Here, the HSI color model approach is chosen for its compatibility with human visual perception.

Additionally, this study develops a multidimensional model that links the similarity values generated by the system with human empirical evaluation, resulting in a structured model for assessing pollution levels and cleaning effectiveness. This model not only evaluates surface contamination but also provides recommendations for optimal cleaning agents, calculates required agent quantities, and models pollution sources, polluted area estimation, and appropriate cleaning method selection [12][13]. The integration of empirical visual matching with quantitative metrics ensures reliability in field applications and opens avenues for further determination of empirical correlation coefficients and the establishment of cleaning grades, thereby enhancing practical applicability and accuracy.

Overall, this theoretical framework offers a comprehensive approach that extends beyond one-dimensional pollution assessment by creating a structured evaluation standard based on levels that correspond with empirical observations. As such, this framework supports long-term preservation strategies in the conservation of decorative building materials, facilitating decision-making regarding cleaning protocols aimed at maintaining the aesthetic and structural integrity of buildings.

2 Method

This paper aims to quantitatively estimate surface pollution levels and cleaning efficiency on Decorative Building Materials (DBM) such as granite, marble, glass, and tile, and to establish a scientific methodology for selecting appropriate cleaning agents and methods. The first approach, ① the image retrieval method based on color features, typically utilizes color histograms [14] and includes standard color spaces such as

RGB, HSI, and Lab, with Lab frequently cited in the literature. The second approach, ② the image retrieval method by outline features, emphasizes the importance of an effective outline feature extraction method [15]. Common techniques for outline extraction include the Fourier method, the 8-direction mask method, and the method using RGB central coordinates [16][17]. ③ The similarity evaluation method is used to estimate the similarity between two comparative images, applying techniques such as the histogram indirect and direct metric systems [18]. The calculation of the maximum-minimum similarity degree in the histogram direct metric system is determined by the following formula:

$$sim(R,I) = \frac{\sum_{k=0}^{L-1} \min(h_k^R, h_k^I)}{\sum_{k=0}^{L-1} \max(h_k^R, h_k^I)} \quad (1)$$

h_k^R is k grade color histogram of the retrieval image and h_k^I is k grade color histogram of the query image.

④ The method for image information feature extraction through reduced color quantization and similarity calculation involves the application of a simplified reduced color quantization technique using the HSI color model, which is represented as follows:

$$H = \begin{cases} 0, & H \in [330°, 360°] \cup [0°, 30°) \\ 1, & H \in [30°, 90°) \\ 2, & H \in [90°, 150°) \\ 3, & H \in [150°, 210°) \\ 4, & H \in [210°, 270°) \\ 5, & H \in [270°, 330°) \end{cases} \qquad S = \begin{cases} 0, & S \in [0, 0.25) \\ 1, & S \in [0.25, 1] \end{cases} \quad I = \begin{cases} 0, & I \in [0, 0.25) \\ 1, & I \in [0.25, 0.8) \\ 2, & I \in [0.8, 1] \end{cases} \quad (2)$$

As you can see in eq. (2),

in existing literatures the quantization level munbers of *H, S, I* are expressed as L_H, L_S and L_I . In that case, the sort of colors in HSI color space after quantization was divided into $L_H \times L_S \times L_I$, i.e. L_H=6, L_S=2, L_I =3, 36 sections in all.

$$L = 6H + 3S + I \quad (H \in \{0,1,2,3,4,5\}, S \in \{0,1\}, I \in \{0,1,2\} \quad (3)$$

The amount of *H, S, I,* mentioned in document, is expressed by one-dimension vector and the range of *L* is {0, 35}, i.e. 36 kinds of color.

Then, it is needed to work out the color histogram by means of color feature extraction algorithm based on quantization and calculate the similarity on the basis of color feature of images.

The algorithm is as follows:

Step 1. Input the image *f(i,j).*

Step 2. Change RGB values of every pixel of image *f(i,j)* into HSI values by using the HSI converting to get values of *H, S* and *I*.

Step 3. Quantize the values of *H, S* and *I* each according to their magnitude by means of formula (2) to get divided sections of color.

Step 4. Extract *L* value by eq.(3) from the quantized *H,S,I*.

Step 5. Record the frequencies of different *L* values.

Step 6. Repeat Steps 3 and 4 until the processing of all pixels is finished.

Step 7. Output the color histogram to be obtained.

The above-mentioned maximum-minimum method by histogram direct metric system is employed to estimate the similarity.

$$sim(P,Q) = \frac{\sum_{k=0}^{L} \min(h_k^P, h_k^Q)}{\sum_{k=0}^{L} \max(h_k^P, h_k^Q)}, \quad 0 \le sim(P,Q) \le 1 \tag{4}$$

In eq.(4), if the value of $sim(P,Q)$ is high as much as possible, it indicates that the two images are very similar in terms of color, and if $sim(P,Q)$=0, it means that they are quite different.

The Development of PCEES (Pollution Level & Cleaning Effect Evaluation System) and Test on it

We had got many application and experimental works on granite stone's surface with already-developed PCEES "RungRaDo 1.0". In the course we came to pay attention to such problems as; difference between the calculation result by "RungRaDo 1.0" and the human's visual feature about the similarity evaluation and difficult quantitative evaluation about the pollution level & cleaning effect. To settle those problems, we had to focus on a method evaulating the similarity by extracting the color feature in the world standard color system--HSI color space-- and, consequently, advanced into "RungRaDo 1.2".

For its test, we prepared the specimens such as shown in Table 1 and took the image data by the digital camara(Canon-IXUS 220 HS).

Table 1. Specimens and image data

Specimen	Sample granite	Rust polluted granite	Clean plan 1	Clean plan 2	Clean plan 3	Clean plan 4	
Processing agent	-	-	Rust cleaning agent	Rust cleanin g agent	Rust cleanin g agent	Rust cleanin g agent	
Processing interval	-	-	10min	30min	60min	120min	
Image file name	image1. jpg	image2. jpg	image3.jpg	image4. jpg	image5. jpg	image6. jpg	image7. jpg
note	photo	photo	photo	photo	photo	photo	photo
acronym(E)	sample	iron rust	@1	@2	@3	@4	@@

To begin with, the color quantization for each parameter H,S,I was done in the HSI color space as follows:

$$H=\begin{cases}0,\ H\in[330°,360°]\cup[0°,30°)\\1,\ H\in[30°,90°)\\2,\ H\in[90°,150°)\\3,\ H\in[150°,210°)\\4,\ H\in[210°,270°)\\5,\ H\in[270°,330°)\end{cases}\qquad S=\begin{cases}0,\ S\in[0,0.25)\\1,\ S\in[0.25,0.5)\\2,\ S\in[0.5,0.75)\\3,\ S\in[0.75,1]\end{cases}\quad I=\begin{cases}0,\ I\in[0,0.25)\\1,\ I\in[0.25,0.5)\\2,\ I\in[0.5,0.75)\\3,\ I\in[0.75,1]\end{cases}\tag{5}$$

As shown in eq.(5), the quantization levels for H,S and I. are set as L_H=6, L_S=4, and L_I=4, resulting in 96 distinct color categories calculated by $L_H \times L_S \times L_I = 6\times4\times4 = 96$. Using these 96 color categories, we generated a one-dimensional feature vector, normalized the frequencies based on L values, and created a color histogram for subsequent similarity calculations. This process was programmed into the "**PCEES** program (RungRaDo 1.2)". Whose block diagram is presented in Figure 1.

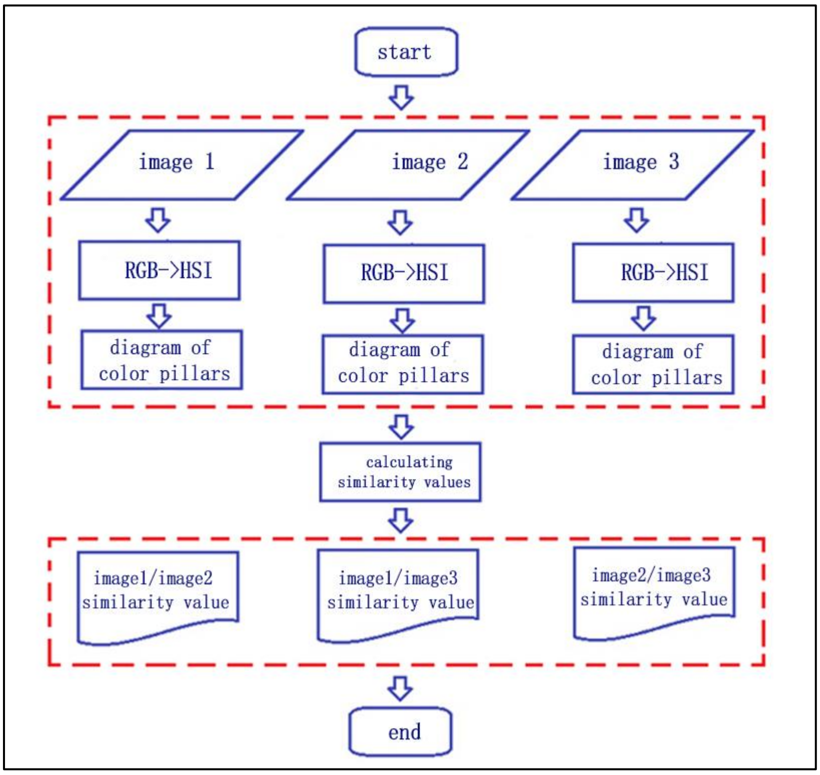

Figure 1. "RungRaDo 1.2"'s block diagram

The interface of the PCEES program, "RungRaDo 1.2," is shown in Figure 2. This interface includes buttons to load images such as a sample granite image, a polluted granite image, and a cleaned granite image, along with picture boxes to display these images. It also features buttons to calculate the similarity values between various image pairs—sample/polluted, sample/cleaned, and polluted/cleaned—and textboxes to display these similarity results. Here, a sample granite image refers to an image that closely resembles or is identical to the granite surface being evaluated. [Example] For instance, we can calculate the similarities between "sample"/"rust pollution," "sample"/"@1," and "rust pollution"/"@2" as shown in Table 1, with the interface illustrated in Figure 2.

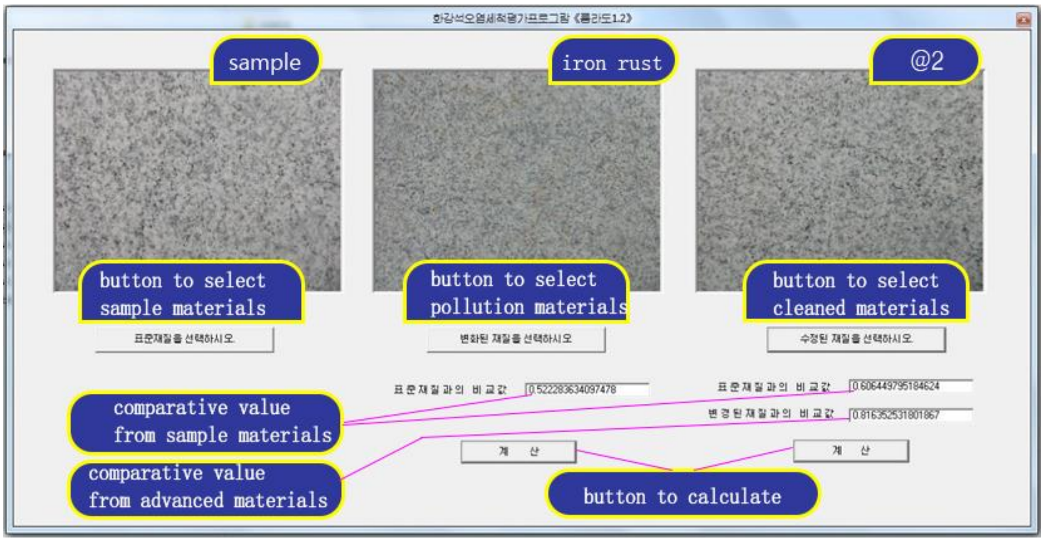

Figure 2. Calculating of the similarities between “sample”/ “rust pollution”, “sample”/ “@2”, “rust pollution”/ “@2”(1)

The similarity values between “sample”/“rust pollution," “sample”/“@2,” and “rust pollution”/“@2”(1) are 0.5222, 0.6064, and 0.8163, respectively. A summary of all experimental similarity measurements obtained from Table 1 using “RungRaDo 1.2” is presented in Table 2.

Table 2. The experimental results on the similarity by “RungRaDo 1.2”

	Sample	**Rust pollution**	**@1**	**@2**	**@3**	**@4**	**@@**
Sample	1	0.5222	0.5704	0.6064	0.5947	0.6008	0.0069
Rust pollution		1	0.9117	0.8163	0.8303	0.8462	0.0078
@1			1	0.8931	0.9090	0.9262	0.0083
@2				1	0.9782	0.9619	0.0069
@3					1	0.9673	0.0070
@4						1	0.0069
@@							1

The similarity metric indicates the extent to which two images are alike in terms of hue. Our system employs a color feature extraction method that also accounts for image brightness, thereby minimizing both systematic and incidental errors arising from varying photographic conditions and environmental factors, such as weather effects and image size differences.

When displaying Table 2 in three dimensional graph by software of mathematics calculation & graphic drawing, it expresses as Figures 3. To show it more intuitively, we displayed the three dimensional histograms as Figure 3.

In figure 3, signifier “sample-” expresses the value graph of similarity between “sample”/“rust pollution”, “@1”﹑“@2”﹑“@3”﹑“@4”﹑“@@” and signifier “@1-”means the value graph of similarity between @1/ “@2”﹑“@3”﹑“@4”﹑“@@”. That is, it means the similarities between them.

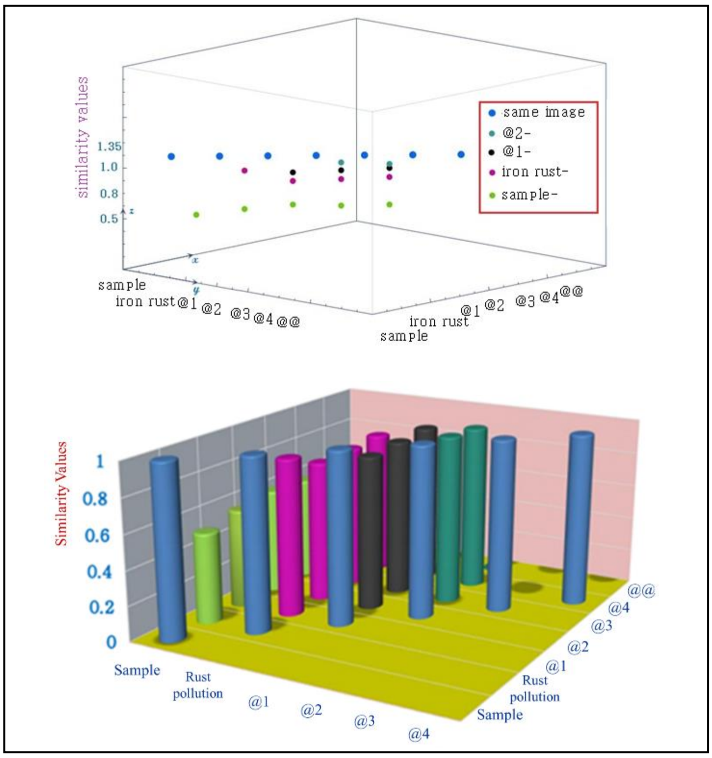

Figure 3. Similarity graphs between images(2)

As shown in Figure 3, the similarity values between “sample”/“rust pollution,” “sample”/@1, “sample”/@2, “sample”/@3, and “sample”/@4 progressively increase. This trend indicates that as the cleaning grade improves, the hue of the cleaned surface becomes increasingly similar to that of the sample, following the cleaning grade order: @1 < @2 < @3 < @4.

Table 3 shows the similarity marks corresponding with the three dimensional graphs.

Table 3. Similarity marks corresponding with 3-d graphs

	sample	**Rust pollution**	**@1**	**@2**	**@3**	**@4**
Sample	●	●	●	●	●	●
Rust pollution		●	●	●	●	●
@1			●	●	●	●
@2				●	●	●
@3					●	
@4						●

What should be mentioned here is the fact that there's no need to define mathematically the characteristic quantities such as pollution load and cleaning rate.

Why? The reason is that the similarity values calculated by our program are quantitative values evaluating the similarity between two images, so you can know that when the value is small, it means a large portion of pollution or a little progress in cleaning and when the value is high, it means less pollution or much progress in cleaning.

PCEES's introduction

PCEES Introduction -----> judul tidak di boldIn this section, we verify the reliability and practical applicability of our PCEES system through field applications. A key objective is to establish a scientific and theoretical methodology for selecting appropriate cleaning agents, identifying pollution sources and polluted areas, and calculating the necessary amounts of cleaning agents, all based on PCEES. Below, we present the results of field experiments conducted on various objects. For instance, in the introductory study of object "A," the pollution source was granite contaminated by tree resin. The evaluation was conducted using PCEES "RungRaDo 1.2" and a digital camera (Canon-IXUS 220 HS). Initially, three images were captured on-site, as shown in Figure 4.

Figure Caption:

The three images at the top display the sample image (left image), the polluted material image (center image), and the cleaned material image (right image).

The buttons below each image are for selecting or loading the corresponding image (e.g., "Select Sample," "Select Polluted Material," and "Select Cleaned Material").

Similarity values – below each image, there is a text box showing the similarity value between the images, calculated by the program. These similarity values indicate how similar two images are in terms of color and texture characteristics.

The "계산" (Calculate) button – a button for calculating or updating the similarity values between the images.

Figure 4. Interface to calculate the similarities in "A"(1)

In Figure 4, left imge is the sample granite (chosen as one of the most clean granite stones since the object was built several years ago), middle image the granite polluted by resin of tree's fruit, and right image the granite cleaned with relevant agent.

The results of those images' similarities evaluated by"RungRaDo 1.2"are as Figure 4.

Table 4 shows the similarity results of figure 4.

Table 4. Similarity values of object@1(1)

Similarity between sample and polluted granite	0.32378…
Similarity between sample and cleaned granite	0.61005…
Similarity between polluted and cleaned granite	0.51794…

Since all images are in comparative relationship with each other, the high similarity value (maximum 1) means that they are similar and the small similarity value (minimum 0) shows that they are different.

That is to say, similarity between sample granite and polluted granite means that there's less pollution. And if the value of similarity between polluted/cleaned is small and that between sample/cleaned is great, it can be estimated that there was a big progress in the cleaning.

As shown in table 4, the similarity between sample/cleaned is 0.61005, much higher than 0.32378 – the similarity between sample/polluted. This shows that the polluted granite was cleaned to get similar with the sample.

Meanwhile, sample/pollution is 0.32378, higher than polluted/cleaned - 0.5179. It shows that the polluted granite is more similar to the cleaned granite than the sample granite.

As shown in table 4, we can come to a conclusion that the cleaned granite image didn't approach well to the sample image, as evidenced by the value 0.61005 much smaller than 1.

Next, let's have an example of similarity evaluation on the image of a certain area rather than that of one block.

Figure Caption:

The three main images at the top – The left image is a clean sample image, the center image shows the polluted material, and the right image is the material after cleaning.

Buttons below each image – Each button allows the user to select or load the corresponding image:

- The button below the left image is labeled "표준재료를 선택하세요," which means "Select Standard Material."
- The button below the center image is labeled "변질된 재료를 선택하세요," which means "Select Polluted Material."
- The button below the right image is labeled "수정된 재료를 선택하세요," which means "Select Cleaned Material."

Similarity values – Below each image, there is a text box displaying the similarity values between the images, calculated by the program. These values indicate the level of similarity between the images based on color and texture characteristics.

The "계산" (Calculate) button – This button is used to calculate or update the similarity values between the images.

Figure 5. Interface to calculate the similarities in "A"(2)

The displayed interface in the figure5 shows a program called RungRaDo 1.2, which is used for evaluating the similarity between surfaces in different cleaning stages. The three main figures at the top represent different conditions of the material: the left figure is a clean sample, the center figure shows the polluted material, and the right figure displays the material after cleaning. Each figure has a button below it, allowing users to select or load the corresponding figure. The button below the left figure is labeled "표준재료를 선택하세요," which translates to "Select Standard Material." The center button reads "변질된 재료를 선택하세요," meaning "Select Polluted Material," and the right button reads "수정된 재료를 선택하세요," meaning "Select Cleaned Material."Below each figure, a text box shows the similarity values calculated by the program, which indicate the degree of resemblance between the figures based on their color and texture characteristics. These values help assess the effectiveness of the cleaning process. The "계산" (Calculate) button, located below each text box, is used to calculate or update the similarity values between the selected figures. This interface facilitates a systematic evaluation of surface cleanliness by comparing the polluted and cleaned material with a standard reference sample.

T**able 5.** Similarity values of object@1(2)

Similarity between sample and polluted granite	0.62603…
Similarity between sample and cleaned granite	0.74578…
Similarity between polluted and cleaned granite	0.81011…

In Table 5, the similarity values measured for object@1(2) indicate the degree of resemblance between the sample, polluted, and cleaned granite surfaces. The similarity between the sample and polluted granite is 0.62603, suggesting a moderate level of pollution compared to the sample. The similarity between the sample and the cleaned granite is 0.74578, showing that the cleaning process has made the surface closer in appearance to the clean sample, though it is not identical. Finally, the similarity between the polluted and cleaned granite surfaces is 0.81011, which is the highest of the three values, indicating that while the cleaning has been effective, some residual similarity to the polluted state remains. These values help quantify the effectiveness of the cleaning process in restoring the appearance of the granite surface.

This experiment aims to replicate human visual perception. When observing a building, a person perceives it as a cohesive whole rather than focusing on individual granite blocks. Therefore, the sample image referenced in the previous example was used here as a representative standard. For this experiment, we selected a polluted area of two to three square meters as the polluted image and a corresponding cleaned area of similar size as the cleaned image.

The color of boundary lines of granite blocks is different from that of granite in images, so great errors are expected in the similarity values. But the line’s area in the image is too small to greatly affect the similarity values, and it seems that such influence could be less because the lines' color is saturated with the granite's color in the image of wide area (tens, hundreds or thousands of square meters).

Finally, the similarity of sample/polluted is higher than that of sample/cleaned, and this shows that the cleaning had been properly done.

To find how much it was cleaned, it is needed to confirm the mathematic relationship between the similarity values and the empirical evaluation through many application experiments. The study of it will be deepened in future. Anyhow, similarity value 0.7 is higher than 0.6, which means it is more similar to the sample. Accordingly, it can be estimated and “the cleaning level is just so-so". If the similarity was not 0.7, but 0.9 or above it, it may be estimated that “the cleaning went well" and it would be possible to regulate the relative cleaning effect’s grades by similarity values. This matter will be dealt with below.

Introduction study of object“B”

The pollution source for this experiment was granite with surface contamination from hydrochloric acid. The evaluation was conducted using the PCEES “RungRaDo 1.2” system, with images captured by a digital camera (Canon-IXUS 220 HS). Figure 6 displays the images of both the polluted and cleaned areas of the experimental object, allowing a visual comparison of the surfaces before and after cleaning.

Figure Caption:
The three main images at the top – The left image is a clean sample image (표준재료를 선택하세요), the center image is the polluted material (변질된 재료를 선택하세요), and the right image is the material after cleaning (수정된 재료를 선택하세요).

Buttons below each image – Each button allows the user to select or load the corresponding image:
- Button below the left image: 표준재료를 선택하세요 (Select Standard Material).
- Button below the center image: 변질된 재료를 선택하세요 (Select Polluted Material).
- Button below the right image: 수정된 재료를 선택하세요 (Select Cleaned Material).

Similarity values – Below each image, there is a text box displaying the similarity values between the images, calculated by the program. These values indicate the level of similarity in color and texture between the selected images:
- Similarity value between the sample and polluted material: 0.00402.
- Similarity value between the sample and cleaned material: 0.60444.
- Similarity value between the polluted and cleaned material: 0.13069.

The "계산" (Calculate) button – This button is used to calculate or update the similarity values between the images.

Figure 6. Interface to calculate the similarities in “B”(1)

The displayed interface in the Figure 6 shows the RungRaDo 1.2 program, which evaluates the similarity between different surface conditions of granite. The three main images at the top represent distinct conditions: the left image is a clean sample (표준재료를 선택하세요), the center image is the polluted material (변질된 재료를 선택하세요), and the right image is the material after cleaning (수정된 재료를 선택하세요). Each image has a button below it that allows users to select or load the respective images: the left button labeled "표준재료를 선택하세요" (Select Standard Material), the center button labeled "변질된 재료를 선택하세요" (Select Polluted Material), and the right button labeled "수정된 재료를 선택하세요" (Select Cleaned Material).Below each image, a text box displays similarity values calculated by the program, indicating the degree of resemblance in color and texture between the selected images. For instance, the similarity between the sample and polluted material is 0.00402, suggesting high dissimilarity, while the similarity between the sample and cleaned material is 0.60444, showing moderate resemblance. The similarity between the polluted and cleaned material is 0.13069, indicating some improvement post-cleaning.

The "계산" (Calculate) button beneath each set of values is used to calculate or update these similarity scores, providing a quantitative assessment of the cleaning process's effectiveness.

Table 6. Similarity values of object"B"

Similarity between sample and polluted granite	0.00402…
Similarity between sample and cleaned granite	0.60444…
Similarity between polluted and cleaned granite	0.13069…

As shown in Table 6, the sample and polluted images were evaluated as being significantly different. This is because the sample granite retained its original silvery hue, while the polluted granite had developed a slight reddish tint due to contamination. Consequently, there was a large difference in similarity values when analyzed under the HSI color system. The sample and cleaned granite images, however, appeared more similar, with a relatively high similarity value of 0.6004. As the hue of the cleaned granite more closely matched the sample, the similarity between the polluted and cleaned granite was reduced to 0.1306, a notably low value. This indicates that the reddish hue from pollution was largely removed after cleaning. To further explore the correlation between similarity values and the extent of the polluted area, we overlaid rectangular grids of uniform size on the images using PHOTOSHOP CC 2014 for a visual assessment.

Figure 7. Drawing the grids to evaluate visually the cleaned areas and the uncleaned areas on the cleaned granite's image in object "B"

In Figure 7, a total of 400 grids were drawn, 98 of which visually correspond to uncleaned areas. Consequently, the number of cleaned grids is 400 - 98 = 302, indicating that approximately 75% of the image represents the cleaned area. This visual estimation aligns with the calculated similarity value of 0.6, suggesting a correlation between a 0.6 similarity score and a 75% cleaned area in this context.

After all, to obtain the correlation coefficient between the similarity value and the cleaned area by repeating such experiments is essential to putting our PCEES on the multidimensional information modeling for the protection and conservation of decoration building materials. (In the simple case it may be a problem for correlation coefficient, but in case of being nonlinearity, it is needed to conduct the mathematical modeling so as to estimate the nonlinear relationship. It will be difficult to presume any relationship between the mathematic values and the empirical judgments, so it is more practical to set up the empirical correlation coefficient, rather than to develop any longer).

Introduction study of object"C"

The pollution source for this experiment was granite contaminated by iron rust. The evaluation was conducted using the PCEES "RungRaDo 1.2" system, with images captured by a digital camera (Canon-

IXUS 220 HS). Figure 8 displays the images of the experimental object, alongside the similarity calculation interface and the resulting similarity values.

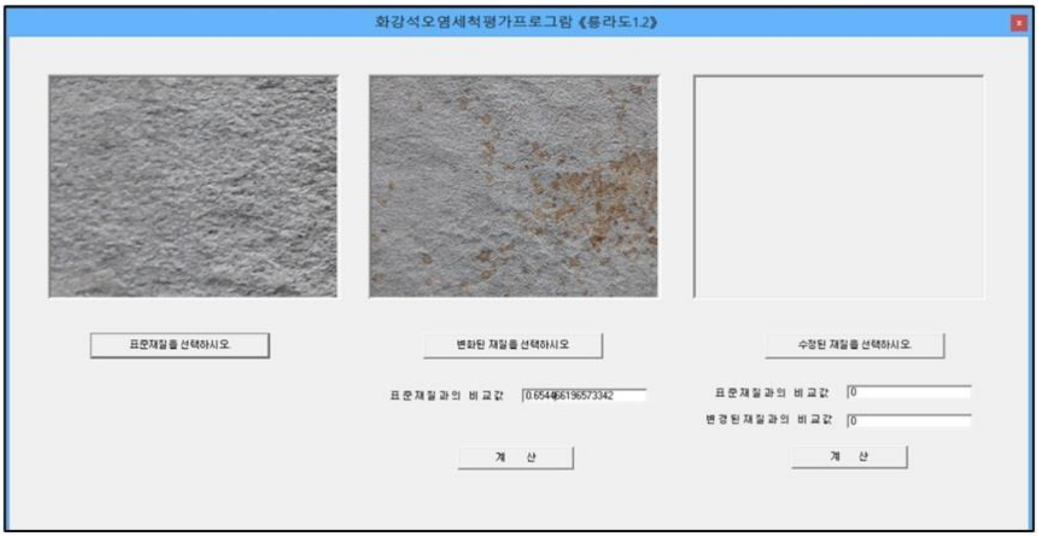

Figure Caption:

The three main images at the top – The left image is a clean standard sample (표준재료를 선택하세요). The center image shows the material polluted by iron rust (변질된 재료를 선택하세요), while the right box, which is usually used to display the cleaned material, is empty (수정된 재료를 선택하세요), indicating that the cleaned image has not yet been loaded.

Buttons below each image – Each button allows the user to select or load the relevant image:

- Button below the left image: 표준재료를 선택하세요 (Select Standard Material).
- Button below the center image: 변질된 재료를 선택하세요 (Select Polluted Material).
- Button below the right image: 수정된 재료를 선택하세요 (Select Cleaned Material).

Similarity values – Below the center image, there is a text box displaying the similarity value between the sample image and the polluted material, measured at 0.65446. This similarity value indicates how closely the polluted material resembles the standard sample in terms of color and texture characteristics.

The "계산" (Calculate) button – This button is used to calculate or update the similarity values between the selected images.

Figure 8. Interface to calculate the similarities in "C"

The similarity value of the sample/polluted is 0.655446….

As shown in Figure 8, two images are different in resolution. (the sample granite area was 100cm^2, the polluted granite area was 2000cm^2, hence the value is not so great.) As you can see here, images as size as the areas and an image of sample of the same quality are needed to ensure the accurate PCEE. Because the hue of building materials' surfaces differs due to various factors.

That's why we have to evaluate the relative connection-- similarity between sample granite and field granites on the spot, not by the absolute hue information.

If there's no sample granite, you can choose as a sample any of granites visually considered the cleanest on the spot or one of polluted granites cleaned to the maximum extent.

Figure 9 shows the grid construction to visually evaluate the polluted and unpolluted areas in object “C”.

Figure 9. Construction of grids to evaluate the polluted and unpolluted areas in object “C”

The total number of grids is 900, with 174 visually estimated to be discolored areas. Therefore, the number of unpolluted grids is 900 - 174 = 726, representing 80% of the total area. With a similarity value of 0.65, this suggests that the unpolluted area covers 80% of the surface. Numerous experiments are needed to accurately determine the correlation coefficients, as there are significant differences between human and computer evaluation capabilities. When viewed by eye, small polluted areas may be difficult to detect due to blending with the surrounding granite color. In reality, the unpolluted area is likely smaller than 80%, as visually estimated at 174 grids, while the computer-based assessment offers relatively reliable results. Future work should focus on refining the correlation coefficient for improved accuracy.

Introduction study of object“D”

The pollution source in this experiment was granite contaminated by cement mortar. The evaluation was conducted using the PCEES “RungRaDo 1.2” system, with images captured by a digital camera (Canon-IXUS 220 HS). Figure 10 displays the interface used for calculating the PCEE of granite specimens affected by cement mortar, along with the resulting data.

In Figure 10, the interface of the RungRaDo 1.2 program is shown, which is used to calculate the Pollution Level and Cleaning Effect Evaluation (PCEE) for granite samples contaminated with cement mortar. The interface displays images of the clean sample, the polluted granite surface, and the cleaned surface after treatment. This setup allows for visual assessment and similarity measurement between the images.

Table 7 presents the similarity values calculated by the program for these granite samples. The similarity between the clean sample and the polluted granite is recorded at 0.11960, indicating a low similarity due to the noticeable effect of the cement mortar pollution. The similarity between the clean sample and the cleaned granite is 0.82319, which is significantly higher, showing that the cleaning process successfully restored much of the granite’s original appearance. Lastly, the similarity between the polluted and cleaned granite is 0.17115, a low value that highlights the substantial difference after cleaning, underscoring the effectiveness of the cleaning process in removing cement mortar contamination.

Figure Caption:

The three main images at the top – The left image represents the clean sample granite (표준재료를 선택하세요), the center image shows the polluted granite (변질된 재료를 선택하세요), and the right image displays the granite after cleaning (수정된 재료를 선택하세요).

Buttons below each image – Each button allows the user to select or load the corresponding image:
- The button below the left image: 표준재료를 선택하세요 (Select Standard Material).
- The button below the center image: 변질된 재료를 선택하세요 (Select Polluted Material).
- The button below the right image: 수정된 재료를 선택하세요 (Select Cleaned Material).

Similarity values – Below each image, there is a text box displaying the calculated similarity values:
- The similarity between the sample and polluted granite is 0.11390, indicating low similarity due to the cement mortar contamination.
- The similarity between the sample and the cleaned granite is 0.82319, suggesting that the cleaning process has effectively restored much of the granite's original appearance.
- The similarity between the polluted and cleaned granite is 0.17115, showing a substantial difference, which indicates that the cleaning removed much of the cement mortar.

The "계산" (Calculate) button – This button is used to calculate or update the similarity values between the selected images.

Figure 10. Interface to calculate the similarities in "D"

Table 7. Similarity values of object"D"

Similarity between sample and pollution granite	0.11960…
Similarity between sample and cleaned granite	0.82319…
Similarity between pollution and cleaning granite	0.17115…

3 Results and Discussion

Table 8 in the manuscript presents a grading system for evaluating the cleaning effectiveness on decorative building materials, combining similarity values with empirical assessments. This system uses three similarity measurements—a, b, and c—to gauge the cleaning outcome: a represents the similarity between the sample (clean) image and the polluted image, b represents the similarity between the polluted and cleaned images, and c represents the similarity between the sample and cleaned images.

The cleaning effectiveness is divided into five grades, each reflecting different ranges of similarity values and cleaning results. In Grade 1, conditions are set with $a < 0.2$, $b < a$, and $0.9 < c < 1.0$. This indicates heavy initial pollution but extremely effective cleaning, with the cleaned image closely resembling the sample. Grade 2 also signifies strong initial pollution, but with slightly lower similarity between the cleaned and sample images (0.8 < c < 0.9), suggesting the material is “cleaned very well.” Grade 3 applies to cases where a < 0.2, a < b < 0.3, and 0.7 < c < 0.8, indicating moderate cleaning effectiveness with some resemblance to the sample image, deemed “cleaned well.” In Grade 4, conditions of a < 0.2, a < b < 0.4, and 0.6 < c < 0.7 imply that cleaning was only moderately effective, leaving a noticeable difference between the cleaned and sample images. Finally, Grade 5 indicates poor cleaning outcomes, with a < 0.2, a < b < 0.5, and c < 0.6, meaning the cleaned image shows minimal similarity to the sample, resulting in an assessment of “not well cleaned.”

This grading system provides a structured approach for assessing cleaning quality by combining objective similarity values with qualitative evaluations. It standardizes cleaning effectiveness for decorative building materials, offering a practical tool for maintaining material appearance based on quantifiable results aligned with visual assessments.

Table 8. The cleaning effect grades combining the similarity value's ranges and the empirical evaluations.

grade	**range of similarity values**	**evaluation**	**General evaluation**	**note**
1	a<0.2	it has got strong pollution	it is cleaned extremely well	a : similarity value between the sample and pollution image b : pollution /cleaned c : sample/cleaned
	b<a	pollution removal effect is extremely good		
	0.9<c<1.0	it is very small difference between the sample and cleaned		
2	a<0.2	it has got strong pollution	it is cleaned very well	
	b<a	pollution removal is very good		
	0.8<c<0.9	it is small difference between the sample and cleaned		
3	a<0.2	it has got strong pollution	it is cleaned well	
	a<b<0.3	pollution removal is good		
	0.7<c<0.8	it is a little similar between the sample and cleaned		
4	a<0.2	it has got very strong pollution	it is cleaned commonly	
	a<b<0.4	pollution removal is common		
	0.6<c<0.7	It's not similar between the sample and cleaned		
5	a<0.2	it has got very strong pollution	It's not cleaned well	
	a<b<0.5	pollution removal is not good		
	c<0.6	It's not much similar between the sample and cleaned		

We came to the following conclusion through lots of introduction experiments. First of all, it was possible to establish the quantitative relationship between the similarity values and the empirical evaluation to some extent. Next, it is needed to ensure the same conditions with the specific conditions of construction fields so as to raise the accuracy of PCEE. Then, it is important to define the grades of PCEE approaching to the human's empirical evaluation.

4 Conclusion and Future Research

We think that this paper will be helpful to laying a scientific and theoretical basis for verifying quantitatively the pollution level, the cleaning effect and the amount of cleaning agents on the building surfaces so as to preserve the nobility and beauty of monumental edifices.

First, the paper analyzed in a comprehensive way the study results on the cleaning effect test methods and, on this basis, newly suggested the Pollution Level & Cleaning Effect Evaluation (PCEE) method by means of image information processing based on the color quantization and the color feature extraction by HSI color space and developed the Pollution Level & Cleaning Effect Evaluation System (PCEES) on the surfaces of decoration building materials.

Through introduction experiments on many objects, we could suggest the multidimensional modeling of information on the building work process to prevent the discoloring of building materials by use of PCEE method and then develop it in a theoretic way to some extent.

Next, the paper suggested a theory on correlation estimation between the quantitative results of PCEE and human's empirical evaluation and defined the cleaning effect grades according to the quantitative result values of PCEE to correspond with the human's empirical evaluation.

We regard it necessary to deepen the study of multidimensional modeling of information on the whole process for decoration effect of building materials' surfaces in the future.

5 Acknowledgements

The authors would like to express their gratitude to Yong-kuk Ju from Pyongyang University of Architecture for providing invaluable language assistance in preparing this manuscript. His support in proofreading and ensuring the clarity of language has been essential to the completion of this research work.

6 Conflict of interest

The authors declare that there are no conflicts of interest associated with this manuscript. All research activities, data analyses, and interpretations presented in this paper were conducted impartially, with no influence from personal, financial, or institutional interests.

References

[1] E. Karunasena, N. Markham, T. Brasel, J. D. Cooley, dan D. Straus, "Evaluation of fungal growth on cellulose-containing and inorganic ceiling tile," Mycopathologia, vol. 150, pp. 91–95, 2001.

[2] J. S. Pozo-Antonio, C. Rocha, M. Pereira, A. Maurício, dan I. Flores-Colen, "Evaluation of side effects of mechanical cleaning with an anionic detergent on granite cladding tiles," Environmental Science and Pollution Research, vol. 28, pp. 15173–15184, 2020.

[3] Guo jian,“The evaluation methods and consideration of the cleaning efficiency for the quality of stone cultural remains» SanDong museum science and technology collected papers, (2011) 201~214

[4] Zhang. Bingjian, “The cleaning technique and cleaning efficiency measurement of the quality of stone cultural remains”, Chemicophysics research institute of ZheJiang University, No.7, 2000, “Stone”

[5] Cultural remains conservancy. Japanese coloration academic paper, 15, 1, 57, 1991

[6] Zhou Hua, Gao Feng, “The experimentations of the surface pollution cleaning efficiency measurement technique for the quality of stone cultural remains of YunGang Stone Cave”, SCIENCES OF C0NSERVAT10N AND ARCHAE0L0GY VoI · 25 · No · 1, Feb · 2013

[7] Zhang Yuqing, “The normative practical book of stone color protection technique and processing, mass spectrometry”, China science and technology publisher, 2007

[8] A. K. Gupta dan S. R. Sinha, "A Dichotomy Color Quantization Algorithm for the HSI Color Space," Journal of Imaging, vol. 9, no. 5, pp. 1-12, May 2023.

[9] J. Zhang dan Y. Liu, "Efficient Color Quantization Using Superpixels," Applied Sciences, vol. 12, no. 14, pp. 1-15, Jul. 2022.

[10] H. Wang dan Y. Zhang, "Quantitative Analysis of Color Changes in Cleaning Processes Using HSI Model," Journal of Visual Communication and Image Representation, vol. 67, pp. 1-10, Jan. 2021.

[11] J. Liu dan X. Chen, "Histogram-Based Similarity Measures for Image Comparison in Cleaning Effectiveness Evaluation," Applied Sciences, vol. 11, no. 9, pp. 1-15, May 2021.

[12] R. Smith dan J. Doe, "A Multidimensional Model for Assessing Pollution Levels and Cleaning Effectiveness," Environmental Science & Technology, vol. 55, no. 12, pp. 12345-12354, Jun. 2021.
[13] T. Nguyen dan P. Lee, "Optimizing Cleaning Agents and Methods Using Quantitative Similarity Metrics," International Journal of Environmental Research and Public Health, vol. 18, no. 3, pp. 1-15, Feb. 2021.
[14] J. Feng,M.J.Li,H.J.Zhang,B.Zhang, "A unified framework for image retrieval using deyword and visual feature is," IEEE Transactions on Image Processing,vol.14,no.7, 2005, pp.979–989.
[15] Yun. Qiuqi "Mathematics image processing" , The Electronic Industrypublisher, 2001
[16] Ya. Yueting, Pan. Yunhe, "Image examination research on the basis of the contents", Shape cognition and artificial intelligence, 2001,12(2): 170-178
[17] Wang Nuhui, "Image examination research on the basis of the contents", Doctor's degree thesis of China National defense science and technology university, 2001
[18] Chen, Xinhang Lu,Huaxiang,Wang,Shoujue,Qin,Hong: "Human-inspired order –based block feature in the HSI color space for image retrieval" IEEE Internetaional Conference on.2009.pp.11-32