Improved visual saliency estimation on manufactured surfaces using high-dynamic reflectance transformation imaging

Marvin Nurit, Gaëtan Le Goïc, Stephane Maniglier, Pierre Jochum, Hermine Chatoux, Alamin Mansouri

Improved visual saliency estimation on manufactured surfaces using High-dynamic Reflectance Transformation Imaging

Nurit Marvin.[a], Le Goïc Gaëtan[a], Maniglier S.[b], Jochum P.[c], Chatoux H.[a], and Mansouri A.[a]

[a]ImViA Laboratory, Université de Bourgogne Franche-comté, Dijon, France
[b]Technical Center for Mechanical Industry (CETIM), Cluses, France
[c]Francéclat, Technical department, Besançon, France

ABSTRACT

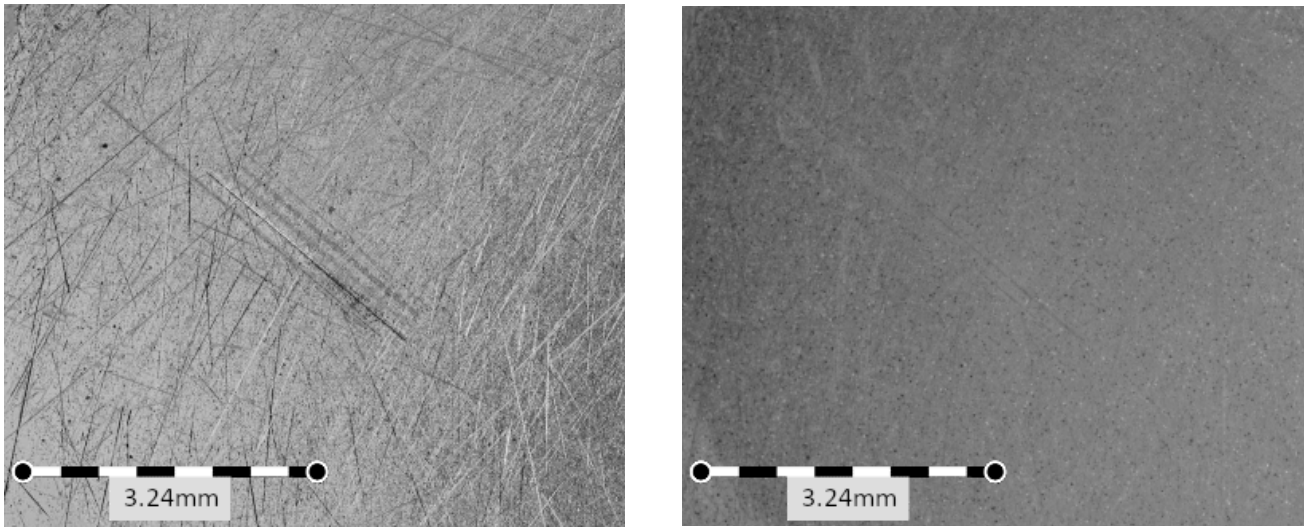
Reflectance Transformation Imaging (RTI) is a technique for estimating the surface local angular reflectance and characterizing the visual properties by varying lighting directions and capturing a set of stereo-photometric images. The proposed method, namely HD-RTI, is based on the coupling of RTI and HDR imaging techniques. The HD-RTI automatically optimizes the necessary exposure times for each angle of illumination by using the response of the scene. Our method is applied to industrial surfaces with micro-scratches from which we will estimate saliency information. Results show that coupling HDR and RTI enhance the characterization and therefore the discrimination on the surfaces visual saliency maps. It leads to an increase in robustness for visual quality assessment tasks.

Keywords: RTI, HDR Imaging, Visual appearance, Visual saliency, Surface quality inspection

1. RTI AND SURFACE QUALITY INSPECTION

Visual inspection of surfaces is a complex sensory process done by human. Therefore, it depends on several factors related to the modalities of the observation: the observer, the object observed, the inspection purpose and the light environment.[1–3] To automate appearance quality inspection, multi-lighting imaging techniques, for example Reflectance Transformation Imaging (RTI), are gaining in interest. The RTI provides access to multidimensional information (dynamic image relighting, luminance, slopes, curvatures, 3D mappings, visual saliency cartographies, *etc.*). These data are very valuable for evaluating and describing the visual quality of a surface,[4–9] especially in the case of manufactured products.[10–15]

The principle of RTI acquisition is to vary the direction of the light source (azimuth and elevation angles) while keeping the sensor fixed orthogonally to the inspected surface for each light position. Thus, in each point (pixel), an evaluation is obtained through a set of discrete luminance values corresponding to the different light directions of the local angular reflectance. The RTI is conventionally implemented with low dynamic range sensors. However, many surfaces require high dynamic due to the materials, their micro-topographies or their heterogeneities. Therefore, capturing angular reflectance often produce a large number of unmeasured, saturated points (under or over-exposed). In addition, the arbitrary lighting adjustment through the sensor exposure time and/or the lighting intensity introduce a potential bias in the RTI data analysis. In figure 1, the local slopes mappings of the same surface are presented for two different exposure time. The perception of geometry is very different while the surface has not changed. It is a limit of the current RTI acquisitions. It is therefore a major issue for the industrial assessment of the surfaces' visual appearance quality, and more globally for industrial inspection. In the specific case of RTI acquisitions, this dynamic limitation occurs at two levels: within the same image (for a particular light position) and for a RTI pixel at each light positions of the RTI acquisition cycle. The proposed approach aims to respond to these two limitations by ensuring that for each illumination angle of the RTI acquisition the number of images and the value of the exposure times necessary for the construction of the HDR information are calculated automatically and adaptively according to the response of the surface. The proposed approach also ensures the fusion of the HDR images constructed for each illumination angle of the RTI acquisition in a common reference frame, in order to allow its analysis and the characterization of the surface behavior. The full dynamic angular reflectance information thus generated is called HD-RTI.[16]



(a) Dy, $T = 125$ ms (b) Dy, $T = 500$ ms

Figure 1. Local slopes mappings obtained from RTI acquisitions of a metallic part with normalized scratches captured at different exposure times (T)

2. METHOD: HD-RTI PIPELINE

The proposed method called HD-RTI consists in coupling the RTI imaging with High Dynamic Range technique (HDR). The coupling is carried out in a self-adaptive manner by automatically estimating the exposure times useful to obtain the full dynamic of the scene, in each angular acquisition position. The number of shots required for each direction of lighting as well as their corresponding exposure times are automatically defined according to the response of the observed surface. The general scheme of the proposed HD-RTI methodology is presented in Figure 2.

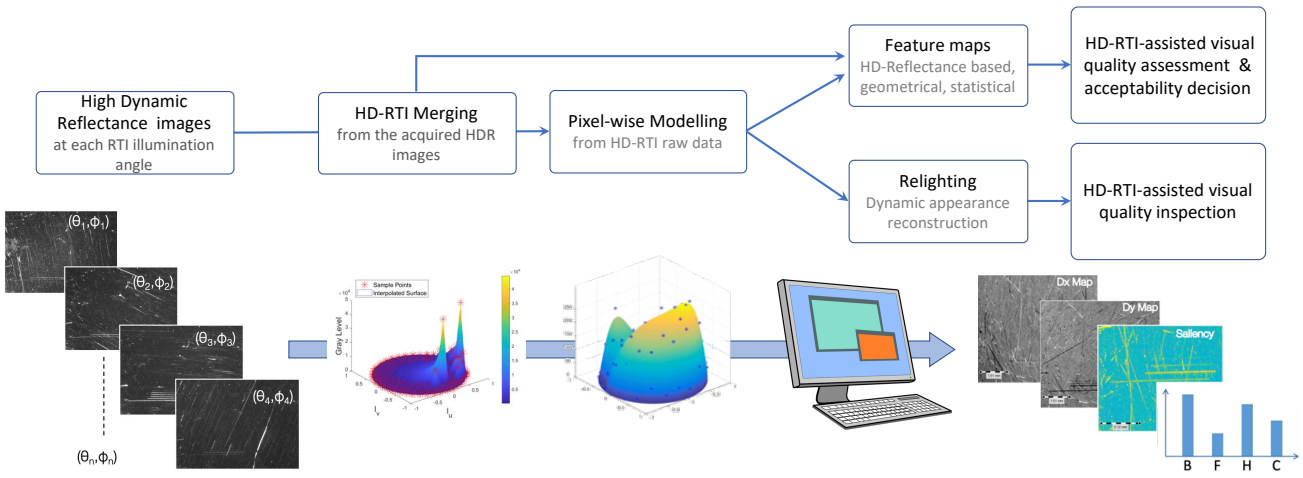


Figure 2. HD-RTI proposed flowchart for quality appearance assessment in industry

This new acquisition technique makes it possible to obtain the information on the local angular reflectance of the surface without the sensor limitation (8 bits) which induces under and over-exposed pixels at a given exposure. In addition, this coupling enable more robust acquisition by removing the bias from the acquisition's

configuration, often done empirically by the operator. Morevover, the descriptors (appearance, geometry or even luminance) extracted from this HD-RTI acquisition benefits from the gained dynamic. This descriptors improvement leads to a more faithful characterization of the captured surface. In an inspection application, we will show that it improves the performance of the salience maps generated from these data and, consequently, the anomalies discrimination.

2.1 HD-RTI methodology

The three steps of the HD-RTI method are detailed below. The first step concerns the acquisition of the high dynamic reflectance images at each illumination angle chosen for the RTI acquisition. This HD acquisition is performed in a self-adaptive way. The second step consist in merging the low dynamic range data acquired into high dynamic range data in order to make the obtained data for all the illumination directions of the acquisition. The third step is associated to the HD-RTI data modeling for relighting.

E1. High dynamic Reflectance Images. The first step of the proposed method consists in implementing an adaptive HDR[17–20] acquisition of the surface reflectance for a light position (θ, ϕ) of the RTI acquisition. The number of shots and the associated exposure times (T) are automatically determined as the acquisition progresses based on the measured data. In practice, the greater the surface reflectance dynamic at a considered illumination angle, the greater the number of necessary images to construct the HDR information, and *vice versa*. Then, we reproduce this method for the other light positions. As the method is self-adaptive, the acquisition time depends on the dynamics of the measured surface, and on the acquisition density chosen in the angular space. However, the current development of sensors which integrate HDR technology and only require a single shot makes it possible to envisage an acquisition time equivalent in the near future to that of a conventional RTI acquisition. For example, for our acquisition device, the acquisition time is increased by a factor of between 2 and 4 according to the surface to be acquired (For the samples in this study the average number of shot is 2.1 with a standard deviation of 0.3).

E2. HD-RTI merging. At this stage, several images $\{I_{T_i}\}_{\theta,\phi}$ are obtained at the different light positions. The i number and T_i values are specific for each light position. The next step is to merge these data together. It is essential for these values to be consistent with one another, and has to be combined to produce a measurement of the surface, contrary to HDR applications for aesthetics purposes. To answer this question, the following method is proposed: a reference exposure time is chosen empirically, denoted T_{ref}, in our case the maximum exposure time. Then, all HDR luminance values are calculated related to this reference exposure time value. The absolute ratio used in the state of the art,[19] induce an intra-image relative reflectance which eliminate the physical link to the surface. Therefore, we chose to rely on the ratio between the exposure time i, denoted T_i and the reference time: $R_T = T_i/T_{ref}$. An HD-RTI acquisition of the inspected surface is obtained, constituted of a set of coherent full dynamic luminance values for each pixel and light position.

E3. Relighting from HD-RTI data. The reconstructed HD-RTI data can be used for appearance reconstruction by virtually varying the amount of light as well as its angle as a function of θ, ϕ and T. This additional degree of freedom for reconstruction allows the lighting configuration to be fully adapted to the area of interest. It is particularly useful for surfaces heterogeneous in color and texture, and/or composite surfaces (multi -materials). To achieve this reconstruction of the dynamic appearance, it is first necessary to build a model from the discrete HDR data acquired at each acquisition angle, making it possible to continuously reconstruct the HDR image of the surface as a function of the angles lighting (θ, ϕ), see Figure 1 step 1. The second step consists in converting the reconstructed HDR image into an LDR image associated with the chosen reconstruction time T, see Figure 1 step 2. This conversion operation is described in Equation 1. The pipeline for the appearance reconstruction with the HD-RTI data is shown in Figure 3.

$$LDR(\theta,\phi,T) = \begin{cases} \text{if } int(HDR(\theta,\phi) \times \dfrac{T}{T_{ref}}) > 255 \text{ then} & 255 \\ \text{else} & int(HDR(\theta,\phi) \times \dfrac{T}{T_{ref}}) \end{cases} \quad (1)$$

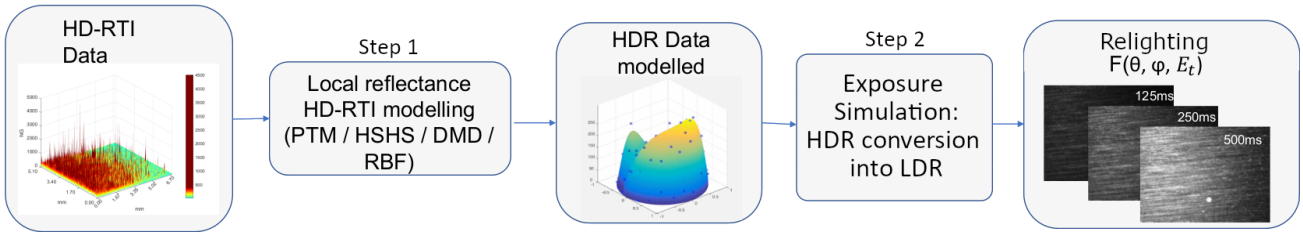


Figure 3. Illustration of the appearance reconstruction from HD-RTI data

2.2 Proof of concept

In order to apply our method of the appearance reconstruction from an HD-RTI acquisition we used a surface sample which is a watch dial that has been preliminarily micro-scratched (Figure 4). The surface is acquired with two acquisitions, the first one, a RTI acquisition ($E_T = 500$ ms) and the second one, a HD-RTI acquisition, both with the same light positions: 149 angular positions homogeneously distributed in an hemisphere above the surface. The appearance was reconstructed with the method of discrete modal decomposition (DMD)[21] at several exposure times from the HD-RTI acquisition and compared with the raw data from the RTI acquisitions which therefore correspond to the ground truth (Figure 5). It can be observed in this figure that the reconstruction of the appearance from HD-RTI data for two particular illumination angles is close to the ground truth and offers in addition, compared with conventional RTI data, the possibility to simulate the exposure time in addition to to the illumination angles the reconstruction. Similar results are obtained for the whole space of directions of illumination. It is thus shown that this approach makes possible to retrieve the conventional RTI data with good performance for any simulated exposure time from an HD-RTI acquisition.

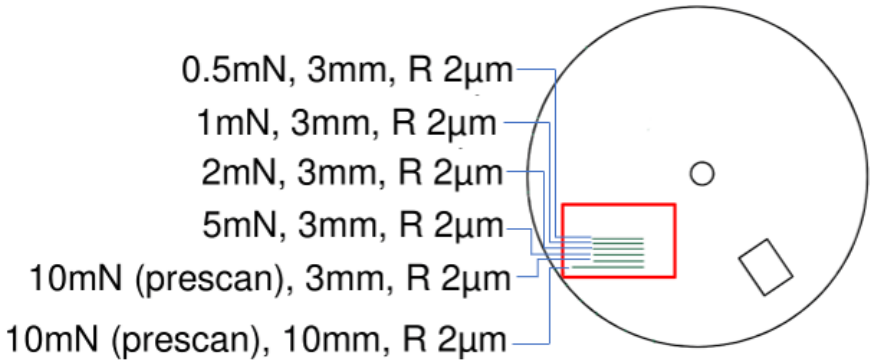
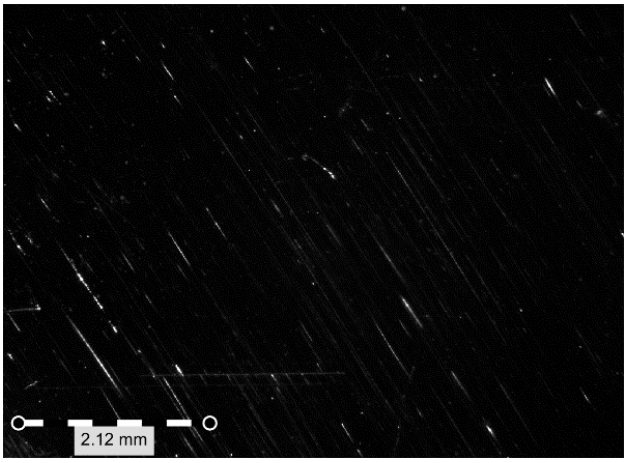
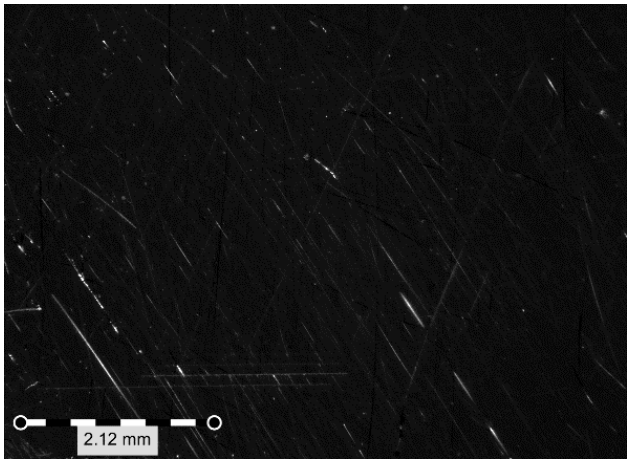


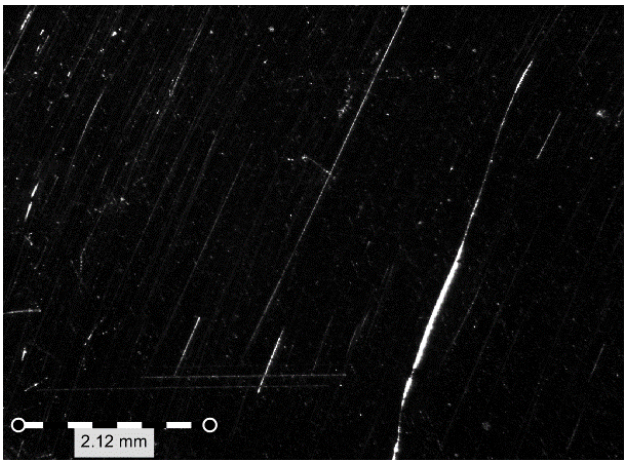
Figure 4. Illustration of the dial and the normalized scratches; The red rectangle indicates the acquired area



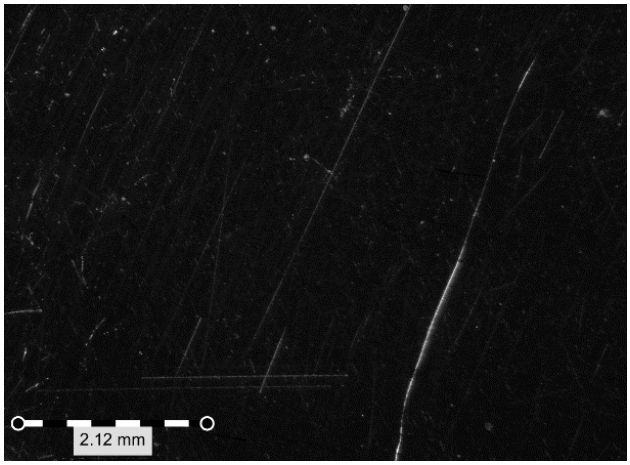
(a) Dial - RTI - Raw Data
$\theta = 197°; \phi = 38°; T = 500$ ms



(b) Dial - HD-RTI - DMD
$\theta = 197°; \phi = 38°; T = 500$ ms



(c) Dial - RTI - Raw Data
$\theta = 308°; \phi = 75°; T = 125$ ms



(d) Dial - HD-RTI - DMD
$\theta = 308°; \phi = 75°; T = 125$ ms

Figure 5. Comparison between the raw data of an RTI acquisition (ground truth) and the appearance reconstruction with the DMD method from an HD-RTI acquisition for different T

3. CHARACTERIZATION AND SALIENCY

To assist the decision-making tasks concerning the appearance quality of manufactured surfaces, the next step is to generate maps derived from an acquisition. To characterize the information,[22] we use descriptors extracted form the HD-RTI acquisition. As indicated in the introduction, the application objectives can be multiple: functional control of the surface states appearance, local defects detection and evaluation and also other non-aesthetic functionalities related to manufactured surfaces. Here we focus on aid for the inspection and assessment of local surface defects. The proposed approach is based on the construction of multi-scale and multi-level saliency maps.[23] The acquired data will be characterized to extract and analyze information about the surface (appearance, texture, anomalies, ...).

From a RTI acquisition of K light positions, with images of $M \times N$ pixels, from each pixel $[P_{(m,n,k)}]_{1 \leqslant k \leqslant K}$, we extract a vector of descriptors $D_{(m,n)}$:

$$P^{RTI}_{(m,n)} = [P_{(m,n,1)}, P_{(m,n,2)}, \ldots, P_{(m,n,K)}], \tag{2}$$

$$\forall P^{RTI}_{(m,n)} \in [1, M] \times [1, N] \; \exists D_{(m,n)} = [d_1, d_2, \ldots, d_L], d_l \in \{\text{descriptors}\}. \tag{3}$$

The descriptors allow to characterize a population in order to determine similarities/differences with another population. In our case, the population is the reflectance values of one pixel, so the characterization is pixelwise. Then from these values, a map can be extracted from a descriptor value at each pixel as illustrated in figure 6.

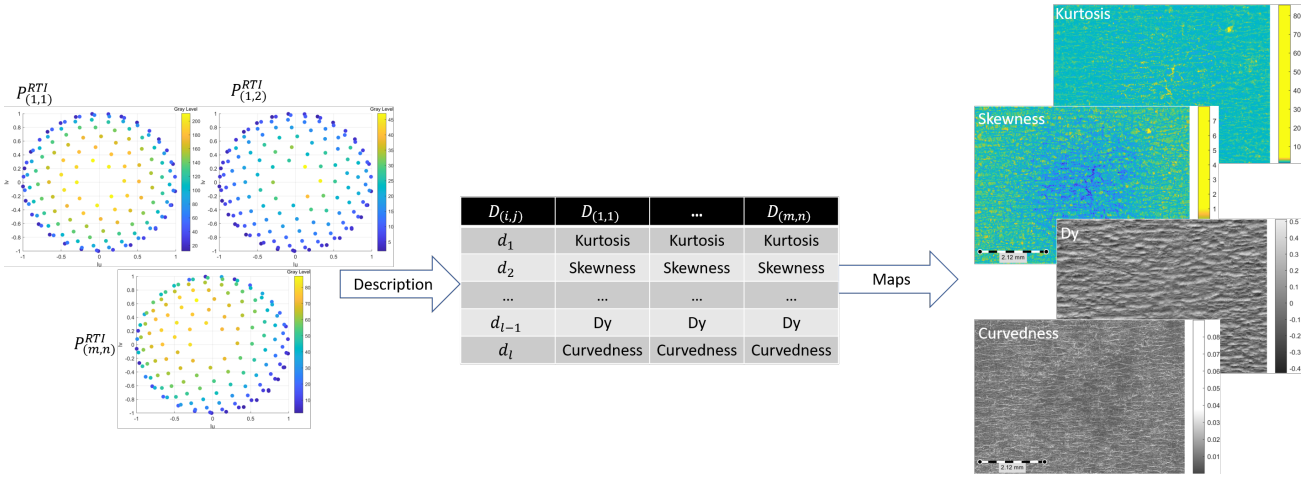


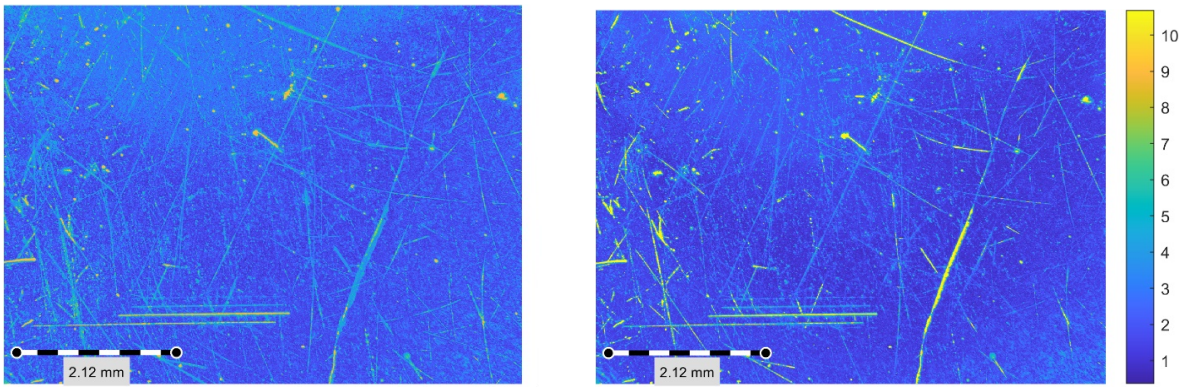
Figure 6. Principle of the characterization of the pixels from a RTI acquisition

The descriptor maps can be used for visual analysis or used to discriminate pixels and enhanced these differences to create a saliency map. The saliency map represents how salient is a pixel by estimating the uniqueness of this pixel among others. The saliency is here estimated from the RTI and HD-RTI descriptors by calculating a Mahalanobis distance[21,24] between a point y and a population X:

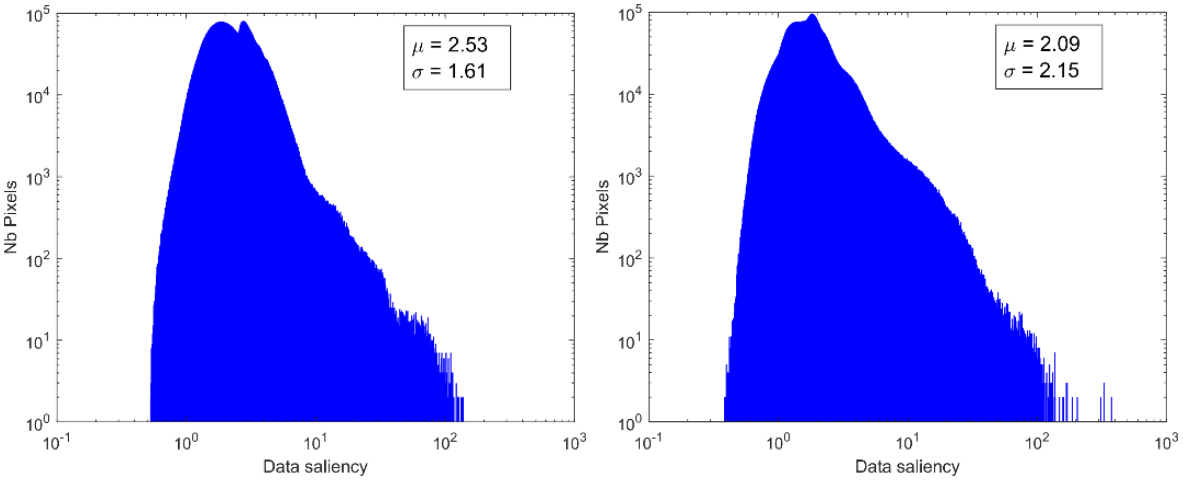
$$\begin{aligned} S^*(y, X) &= \sqrt{[y - \mu(X)]^T \Sigma(X)^{-1} [y - \mu(X)]}, \\ S^*(D_{(m,n)}, D) &\Rightarrow \text{a pixel saliency}, \\ S^*(D, D) &\Rightarrow \text{Pixels saliency map}, \end{aligned} \tag{4}$$

Where Σ is the covariance matrix and μ the mean of the population X.

In the case of this application, the implemented descriptors are estimated from the luminance distributions of each pixel (mean, median, coefficient of variation, kurtosis, skewness, entropy, energy, max and min). In figure 7, the saliency maps are assessed for the two acquisitions described in subsec. 2.2. One can observe in two way that our proposed method generates a better saliency map. Firstly, visually the HD-RTI saliency map highlights anomalies better by enhancing the contrast between the background (frequent behaviors) and the anomalies (unique behaviors). Secondly, this visual analysis can be confirmed by computing the mean and the standard deviation of the data histogram. For our method, the mean is lower and standard deviation greater because the frequent behaviors have lower saliency values, and by their quantities, they decrease the global mean, conversely the unique behaviors have greater saliency values, so the global standard deviation is increased.



(a) RTI derived saliency map (b) HD-RTI derived saliency map



(c) Saliency histogram (RTI) (d) Saliency histogram (HD-RTI)

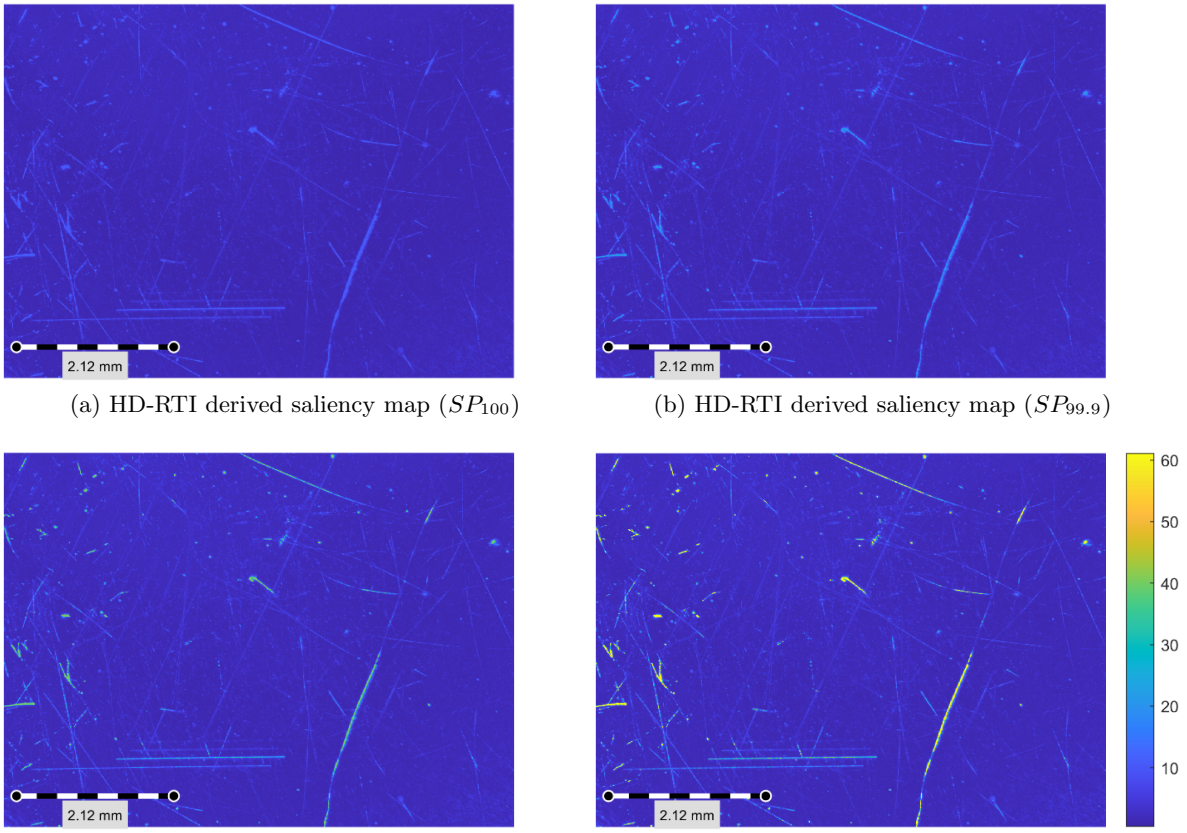
Figure 7. RTI *vs* HD-RTI - Saliency maps and histogram

The saliency of a pixel varies with the chosen neighbors: it increases if its characters are different from those of its neighborhood, and vice versa. When the most unique individuals are removed from the neighbors, only individuals with low uniqueness remain, therefore, the pixels, which have low saliency, will have decreased saliency because they will be compared with similar individuals. Moreover, the pixels, which have high saliency, will have increased saliency because they will be compared with different individuals. To perform this elimination of salient points from the saliency calculation, a first saliency calculation is made on all pixels, then a second saliency is calculated by considering only the pixels with a saliency level below a threshold expressed in cumulative probability (Equation 5).

$$\begin{aligned} \{X^{SP_z}\} &= \{X \quad |P(Sj(X,X)) \leqslant z, z \in [0,1]\} \\ S^{SP_z}(V, V^{SP_z}) &\Rightarrow \text{Pixels multi-level saliency } (\leqslant z) \end{aligned} \tag{5}$$

Where P is the cumulative probability ans z a chosen value.

It is observed in figure 8 a significative increase visual contrast between the non-salient points and those salient (anomalies) with the decrease of the SP value. This method makes it possible to control the threshold value and therefore to the contrast increase between the salient and less salient points, thus improving the discrimination between all the pixels but also between the anomalies, and making the anomalies classification easier.



(a) HD-RTI derived saliency map (SP_{100})

(b) HD-RTI derived saliency map ($SP_{99.9}$)

(c) HD-RTI derived saliency map ($SP_{99.5}$)

(d) HD-RTI derived saliency map (SP_{99})

Figure 8. HD-RTI multi-level saliency maps with different values of SP

Then, another method allow us to modify the phenomena used during the saliency calculation. In fact, saliency is related in terms of characters, but also in terms of observation scale. A phenomenon, compared to its close neighborhood, will seem unique but if we compare it to all the measured phenomena then its characteristics will be more common. The observation scale therefore changes the saliency phenomena estimation. To change our observation scale, we calculate saliency by using a sliding window to limit the pixels population, as described in Equation 6.

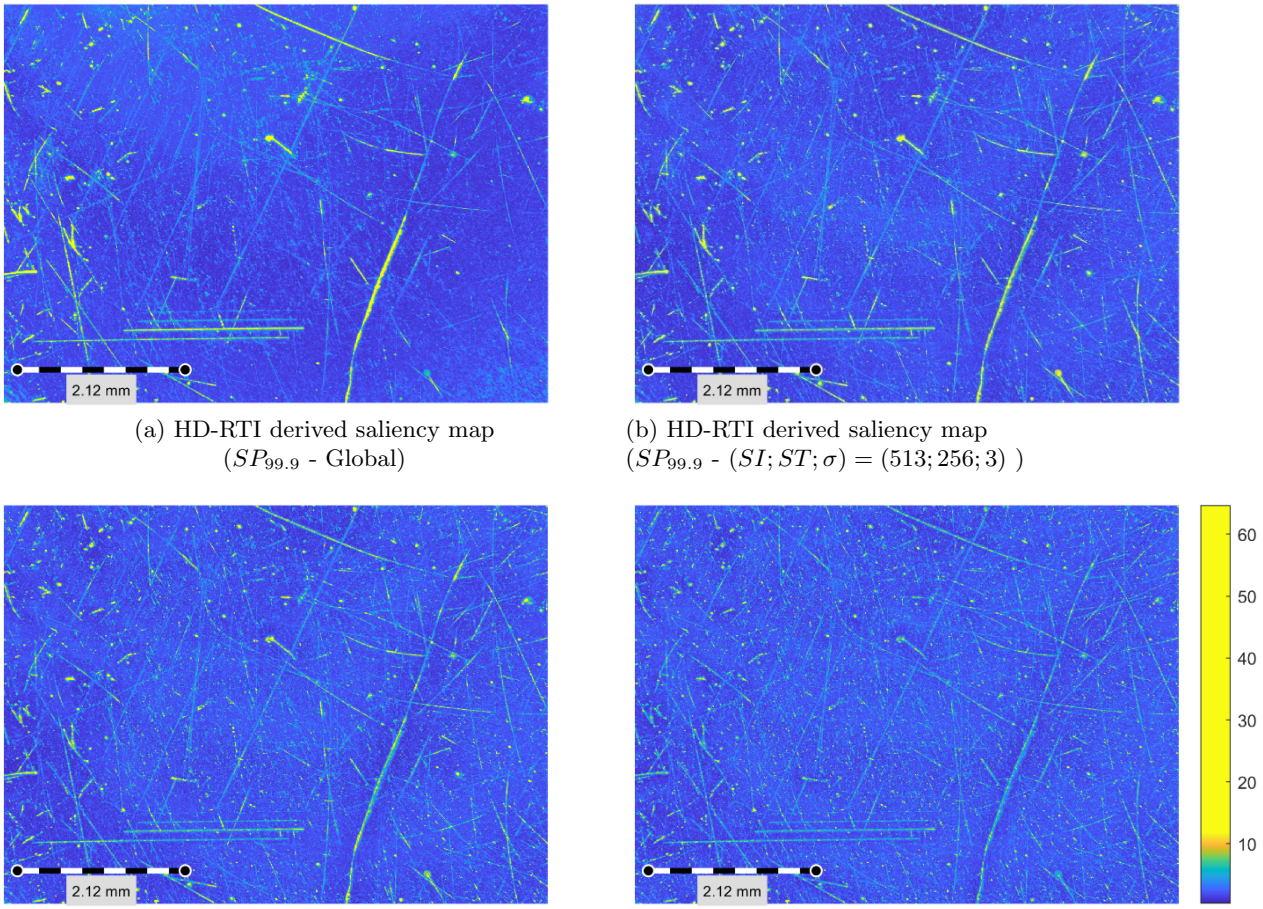
$$S^*_{(i,j)}(D_r, D_r), \tag{6}$$

Where D_r is a window of size $2r$ centered in (i, j) to calculate the pixels saliency $S_{(i,j)}$ on the sliding window.

We can also use equations 5 and 6 to estimate a saliency value at different scales while extracting the excessively salient points from computation (Equation 7).

$$S^{SP_z}_{(i,j)}(D_r, D^{SP_z}_r). \tag{7}$$

However, the saliency of pixels far from the center of the sliding window is less reliable because these pixels have fewer close neighbors. It is therefore important to associate a weight during the calculation at each point, in our case the weight function used is a 2-dimensional Gaussian. There are then 3 parameters to tune the multiscale salience, the size of the sliding window (SI), the step of the translation (ST) and the standard deviation of the Gaussian (STD). We can see the result of the multiscale saliency in Figure 9.



(a) HD-RTI derived saliency map ($SP_{99.9}$ - Global)

(b) HD-RTI derived saliency map ($SP_{99.9}$ - $(SI; ST; \sigma) = (513; 256; 3)$)

(c) HD-RTI derived saliency map ($SP_{99.9}$; - $(SI; ST; \sigma) = (257; 128; 3)$)

(d) HD-RTI derived saliency map ($SP_{99.9}$; - $(SI; ST; \sigma) = (129; 64; 3)$)

Figure 9. HD-RTI multiscale saliency maps at different scales

Figure 9 presents a result of the multiscale saliency. The saliency at different scales, unlike multi-level saliency, increases the saliency of phenomena whose uniqueness is much smaller but still significant to be analyzed. Several anomalies appear because their local saliency are enhanced only by using the close neighborhood. Finally the coupling of the two calculation methods makes it possible to keep the dynamics obtained by the multi-scale but also the contrast increase induced by the threshold.

4. CONCLUSION

This paper presented a methodology to implement High Dynamic Range technique for RTI acquisition. The relevance of the HD-RTI acquisition is investigated by comparing saliency results obtained with RTI and HD-RTI acquisitions. It appears to provide a robust and reliable estimation of the local angular reflectance in terms of luminance. It is particularly relevant to describe the local angular reflectance surfaces in a surface defects detection/characterization context. A perspective of this work is to specialize the descriptors to specific analysis purpose. In the same time, adapting these descriptors to the surface type could improve the saliency. Another perspective is to obtain a multi-scales saliency map by merging saliency at different observation scales. Our objective is to enhance the saliency and the quality assessments task robustness. Finally, a next step in this study will consist in quantifying the performance gain of the proposed approach in terms of visual inspection through psychometric experimentation.[25]

ACKNOWLEDGMENTS

This work benefited of the funding of French National Research Agency (ANR) through the project NAPS (https://anr.fr/Projet-ANR-17-CE10-0005).

REFERENCES

[1] Baudet, N., Pillet, M., and Maire, J. L., "Visual inspection of products: a comparison of the methods used to evaluate surface anomalies," *International Journal of Metrology and Quality Engineering* (2011).

[2] Guerra, A. S., *Métrologie sensorielle dans le cadre du contrôle qualité visuel*, PhD thesis, Université de Savoie (2008).

[3] Debrosse, T., Pillet, M., Maire, J. L., and Baudet, N., "Sensory perception of surfaces quality - industrial practices and prospects," *the Proceedings of the Kansei Engineering and Emotion Research* (2010).

[4] Lewis, D. A., Nurit, M., Chatoux, H., Meriaudeau, F., and Mansouri, A., "An automated adaptive focus pipeline for reflectance transformation imaging," *Electronic Imaging* (2021).

[5] Castro, Y., Pitard, G., Zendagui, A., and Le Goïc, G., "Light spatial distribution calibration based on local density estimation for reflectance transformation imaging," *Proc. SPIE QCAV, Quality Control by Artificial Vision* (2019).

[6] Castro, Y., Pitard, G., Le Goïc, G., and Brost, V., "A new method for calibration of the spatial distribution of light positions in free-form rti acquisitions," (2019).

[7] Castro, Y., Nurit, M., Pitard, G., and Zendagui, A., "Calibration of spatial distribution of light sources in reflectance transformation imaging based on adaptive local density estimation," *Journal of Electronic Imaging* (2020).

[8] Morita, M. M., Novoa, F. D., and Bilmes, G. M., "Reflectance transformation imaging. first applications in cultural heritage in argentina," *Journal of Archaeological Science: Reports* (2019).

[9] Lemesle, J., Robache, F., Le Goic, G., and Mansouri, A., "Surface reflectance: An optical method for multiscale curvature characterization of wear on ceramic–metal composites," *Materials* (2020).

[10] Zendagui, A., Thomas, J.-B., Le Goic, G., Castro, Y., Nurit, M., Mansouri, A., and Pedersen, M., "Quality Assessment of Reconstruction and Relighting from RTI Images: Application to Manufactured Surfaces," *2019 15th International Conference on Signal-Image Technology Internet-Based Systems (SITIS)* (2019).

[11] Martlin, B. and Rando, C., [*An assessment of the reliability of cut surface characteristics to distinguish between hand-powered reciprocating saw blades in cases of experimental dismemberment*], Journal of Forensic Sciences (2020).

[12] Pernkopf, F., "3d surface acquisition and reconstruction for inspection of raw steel products," *Computers in Industry* (2005).

[13] Iglesias, C., Martínez, J., and Taboada, J., "Automated vision system for quality inspection of slate slabs," *Computers in Industry* (2018).

[14] Smith, M. and Stamp, R., "Automated inspection of textured ceramic tiles," *Computers in Industry* (2000).

[15] Coules, H., Orrock, P., and Seow, C., "Reflectance transformation imaging as a tool for engineering failure analysis," *Engineering Failure Analysis* (2019).

[16] Nurit, M., Castro, Y., Zendagui, A., Goïc, G. L., Favreliere, H., and Mansouri, A., "High dynamic range reflectance transformation imaging: an adaptive multi-light approach for visual surface quality assessment," *Proc. SPIE QCAV, Quality Control by Artificial Vision* (2019).

[17] Martínez, M. A., Valero, E. M., and Hernández-Andrés, J., "Adaptive exposure estimation for high dynamic range imaging applied to natural scenes and daylight skies," *Applied Optics* (2015).

[18] Banterle, F., Artusi, A., Debattista, K., and Chalmers, A., [*Advanced high dynamic range imaging*], 2nd Edition, A K Peters/CRC Press (2017).

[19] Debevec, P. E. and Malik, J., "Recovering high dynamic range radiance maps from photographs," *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques* , ACM Press/Addison-Wesley Publishing Co. (2005).

[20] Robertson, M., Borman, S., and Stevenson, R., "Dynamic range improvement through multiple exposures," *Proc. IEEE ICIP, International Conference on Image Processing* (2000).

[21] Pitard, G., Le Goïc, G., Mansouri, A., Favreliere, H., Désage, S.-F., Samper, S., and Pillet, M., "Discrete modal decomposition: a new approach for the reflectance modeling and rendering of real surfaces," *Machine Vision and Applications* (2017).

[22] Brown, C. A., Hansen, H. N., Jiang, X. J., Blateyron, F., Berglund, J., Senin, N., Bartkowiak, T., Dixon, B., Le Goïc, G., Quinsat, Y., Stemp, W. J., Thompson, M. K., Ungar, P. S., and Zahouani, E. H., "Multiscale analyses and characterizations of surface topographies," *CIRP Annals* (2018).

[23] Anzid, H., Le Goic, G., Bekkari, A., and Mansouri, A., [*Benchmarking Saliency Detection Methods on Multimodal Image Data*], Image and Signal Processing (2018).

[24] Mahalanobis, P. C., "On the generalized distance in statistics," *Proceedings of the National Institute of Sciences (Calcutta)* (1936).

[25] Zendagui, A., Le Goïc, G., Chatoux, H., Thomas, J.-B., Pierre, J., Stephane, M., Castro, Y., Nurit, M., and Mansouri, A., "Quality assessment of dynamic virtual relighting from rti data: application to the inspection of engineering surfaces," *Proceedings of the International Conference on Quality Control by Artificial Vision (QCAV 2021)* (2021).