Journal of Visual Communication and Image Representation (2021)

Contents lists available at ScienceDirect

Journal of Visual Communication and Image Representation

journal homepage: www.elsevier.com/locate/jvci

# A-contrario framework for detection of alterations in varnished surfaces

Alireza Rezaei[a,*], Sylvie Le Hégarat-Mascle[a], Emanuel Aldea[a], Piercarlo Dondi[b,d], Marco Malagodi[c,d]

[a] SATIE Laboratory, University Paris Saclay, rue Noetzlin, Gif-sur-Yvette 91190, France
[b] Department of Electrical, Computer and Biomedical Engineering, University of Pavia, Via Ferrata 5, 27100, Pavia, Italy
[c] Department of Musicology and Cultural Heritage, University of Pavia, Corso Garibaldi 178, 26100, Cremona, Italy
[d] CISRiC - Arvedi Laboratory of Non-Invasive Diagnostics, University of Pavia, Via Bell'Aspa 3, 26100, Cremona, Italy

ARTICLE INFO

*Article history*:
Received -
Received in final form -
Accepted -
Available online -

Communicated by -

*2000 MSC:* 68U10, 68T10

*Keywords:* A-contrario framework, Defect detection, Preventive conservation, Historical violins

ABSTRACT

Preventive conservation is the constant monitoring of the state of conservation of an artwork to reduce the risk of damages and so to minimize the necessity of restorations. Many methods have been proposed during time, generally including a mix of different analytical techniques. In this work, we present a probabilistic approach based on the a-contrario framework for the detection of alterations on varnished surfaces, in particular those of historical musical instruments. Our method is a one step Number of False Alarms (NFA) clustering solution which considers simultaneously gray-level and spatial density information in a single background model. The proposed approach is robust to noise and avoids parameter tuning as well as any assumption about the shape and size of the worn-out areas. Tests have been conducted on UV induced fluorescence (UVIFL) image sequences included in the "Violins UVIFL imagery" dataset. UVIFL photography is a well known diagnostic technique used to see details of a surface not perceivable with visible light. The obtained results prove the capability of the algorithm to properly detect the altered regions. Comparisons with other the state-of-the-art clustering methods show improvement in both precision and recall.



## 1. Introduction

Preventive conservation, a crucial procedure in Cultural Heritage, consists in the constant monitoring of artworks and monuments, performed with the aim of mitigating the risk of damages and of minimize restorations [1, 2].

Historical wood musical instruments (such as violins or violas) are a peculiar kind of artworks since they are both held in museums and played (even nowadays), leading to a major risk of mechanical wear and entailing the need of routine monitoring through a non-invasive technique such as optical imaging. Nevertheless this latter presents various complexities: (i) the instruments have undergone multiple restorations during centuries, leading to a stratified surface difficult to analyze; (ii) varnish wear can evolve in different ways depending on the initial conditions of the surface and on the different substances being present; (iii) the rounded morphology and the high reflectance of the varnishes generate noisy reflections during photo acquisition (almost impossible to avoid completely) that can be confused with potential alterations; (iv) in order to avoid damages to the varnishes, violins cannot be rigidly fixed to a support, thus, slight misalignment between photographic sessions can occur. Then, the objective of regular analysis of images is to quickly identify potentially altered areas despite noise, and then, to apply only on them spectroscopic techniques, such as X-Ray Fluorescence (XRF) or Fourier Transform Infrared (FTIR) spectroscopy, as confirmation.. This approach is less time consuming and more focused with respect to traditional monitoring solutions that involve the use of multiple analytical techniques on a large portion of the surface (e.g., continuously

*Corresponding author. E-mail: alireza.rezaei@universite-paris-saclay.fr

check the entire bottom and top areas of the violin back plate) [3, 4].

Wear detection can be viewed as a semantic segmentation problem with two semantic classes, namely the altered region(s) and the unchanged areas. Having a reference image is important since we want to focus on the presence of new alterations. Historical violins have been played for centuries before to be held in museums, thus, they can already have some worn-out regions. Using a single image it is possible to properly identify established wear [5], but, without a reference image, small new alterations would be missed, especially in their initial stage when they can be confused with nuances of the varnishes. Thus, instead of an anomaly detection problem, we would rather cast wear detection as a change detection problem assuming we have at our disposal a reference image representing the initial state of the instrument.

In terms of image processing techniques, various approaches have been proposed to detect change areas that are highly variable both in terms of radiometry and in terms of geometric shape. These approaches include the statistical rejection tests in which decisions are taken based on the probability of the null hypothesis representing the absence of change, binary segmentation including image thresholding techniques and recent machine learning methods, and clustering approaches. However, in the case of tenuous changes due to wear, possibly buried in noise, one has to rely on both spectral and spatial cues to decide on the presence of a change.

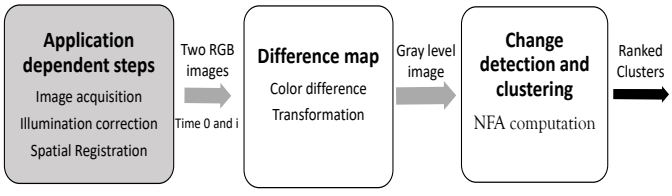



Fig. 1: General steps of the proposed algorithm.

In this work, motivated by our problem of subtle change detection, we define a new method based on the a-contrario framework [6, 7]. The main contribution of the proposed approach is to encompass in a single model both spatial and radiometric criteria that change areas present. Moreover, the significance based decision process allows us to be free from any threshold or parameters characterizing the clusters (shape, number, position etc.). The general layout of our approach is presented in Fig. 1. The first two blocks (discussed in Section 4) deal mainly with sensor data acquisition and computing a raw difference map. Their purpose is basically to construct a normalized input for the change detection, but our core contribution related to the significance based decision corresponds to the third block. The proposed method (which extends [8] both from theoretical and experimental point-of-view) has been validated for early detection of superficial alterations on historical violins. However, we are confident that, by adapting the acquisition and preprocessing steps, a wider range of change detection problems may be addressed as long as the decision is driven by spatial and radiometric considerations at the same time.

## 2. Related works

In recent years, many solutions for the problem of change detection have gone towards the use of deep neural networks. Various fields such as medical imaging and remote sensing have defined several network architectures to address the inherent difficulties of the problem [9, 10, 11, 12, 13, 14, 15]. The fine-tuning process in these works require enough amount of samples to work properly. For example, the network proposed in [16] uses $10^4$ labeled images only for network fine tuning and for the detection of defects in radiography images. However, in optical monitoring for cultural heritage, the acquisition of new data is a time consuming process and expert annotations are not always as precise and consistent as desired.

Working in an unsupervised way, clustering approaches have been proposed to cope with the lack of labeled data. Furthermore, when performed in an adequate feature space, some of them have the ability to consider both spatial and spectral features. However, each family of clustering methods introduces its own driving set of sensitive parameters (cluster number for partition based methods [17], distance threshold for density based methods [18, 19] including hierarchical ones [20], etc.) to be tuned which in our case can be difficult since we do not have any prior knowledge about the number of changed areas (if any) and their relative extent.

The a-contrario framework introduced by Desolneux et al. [6] has proven its efficiency in image analysis in a variety of applications such as texture analysis [21], motion detection [22, 23], edge and line detection [24, 25, 26, 27], or reconstruction from motion [28]. In all these studies, the detection is performed by rejecting a *naive model* which describes the statistic of the unstructured data. At the same time, grouping principles have been used for tasks related to the higher level perceptual organization of scenes [29, 30], thanks to their general nature.

In our case, we are interested in the applications dealing with the detection of changed areas across multi-temporal images. In early studies, the a-contrario framework and spectral invariant features have been used to detect meaningful changes between two satellite images of the same area taken at different times [31]. An a-contrario approach has also been proposed for change detection in three dimensional multi-modal medical images such as Magnetic Resonance sequences [32]. In 2010, Robin et al. [33] use the a-contrario framework for the definition of a criterion assessing the level of coherence in a sequence of images for detecting sub-pixel changes in a time-series of satellite images. Flenner et al. [34] further investigated this approach by using exchangeable random variables instead of relying on the independent and identically distributed (IID) assumption. Previously cited works focused on the gray level values (and their changes) so that the considered naive model deals with gray level discrepancy.

Now, other works have been proposed to deal with objects characterized by their spatial feature consistency. Searching for

a pattern using a-contrario framework can be done either by computing the significance of every pattern corresponding to a given parametric shape (e.g., circle [35], line, ellipse [36]), or even several shape categories at once [37], or by checking possible clusters of points regardless of their shape [38, 39]. In these works, even though one may consider more complex models which allow for taking into account dependencies in the image [40], the usual naive model that represents unstructured data is the uniform distribution, for instance that of the gradient orientations in [41] or simply of the location of the 1-valued pixels in [6] and [42] that process binary images.

Based on all the previously cited works, a straightforward solution for involving gray-level and spatial features considered in [26, 43] is to first use an a-contrario approach to detect points that, according to their gray-level values, are likely to belong to a change area, to store these points as a binary image, and then group these points (1-valued in the binary image) together as clusters when applicable [7]. However, with such an approach the two criteria on gray-level values and spatial features of the researched objects or areas are considered sequentially and there is an implicit threshold (at the end of the first step) that leads to a loss of information regarding the intensity of the change. Our proposal is thus to combine both steps and cluster a gray-level image by using a single naive model. The core idea is the same as the one proposed in [8], but the algorithm itself has been improved to make it more consistent mathematically while some implementation guidelines are provided in order to facilitate the adoption of the underlying idea in broader applications. On the experimental side, we add a study on synthetic data which strengthens the assumptions arising from the a-contrario model, along with tests on more real data and comparisons with other approaches.

## 3. Proposed approach

### *3.1. Notations and key idea*

Although the violin monitoring process collects a series of images acquired at pre-defined time intervals, in the proposed approach, only the current and the reference images are considered. They are denoted $I$ and $I_0$ respectively and, being in conventional RGB format, their comparison by using a color difference formula [44] provides a difference map $\Delta I$ such that $\Delta I$ is a gray-level image defined on the pixel domain $\mathcal{P} \subset \mathbb{N}^2$.

Our problem boils down to segmenting $\Delta I$, with respect to semantic classes, one of which representing the unchanged area. As previously stated, we propose to rely on a single naive model which will account for both radiometric and spatial criteria characterizing a wear area that is present in $\Delta I_i$ images.

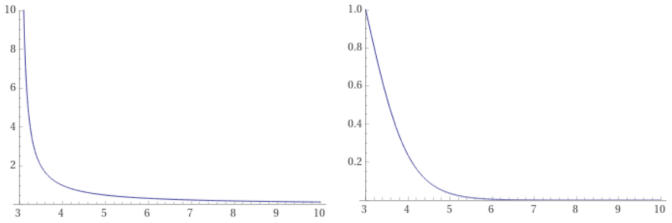
The basic idea is to extend the meaningfulness concept specifying that a cluster is all the more significant that it is very dense (i.e., its points are 'surprisingly' close) not only spatially but also in terms of gray-level differences. Now, to include gray-level features in a-contrario detection, we could either adapt the naive model to gray-level features in case of unstructured data (no change in our case), or adapt the gray-level values so that the uniform distribution can be used as naive model as usual. In this work, we adopt the second approach. Considering gray-level differences, low values correspond (mainly) to no change and high values (mainly) to changes, so that a gray-level transform is required to meet the assumption that a change can be detected as surprisingly structured or dense. Then, using the cluster NFA based on distance, the proposed method also needs the specification of the considered distance. These two points are presented in the next subsections before the presentation of the NFA computation and cluster detection algorithm.

### *3.2. Gray-level transformation*

Let us first enumerate the desirable properties of the required gray-level transformation for $\Delta I$ pixel values: following the transformation, (i) the gray-level values of pixels belonging to unchanged areas shall be stretched, (ii) the gray-level values of pixels belonging to change areas shall be similar and (iii) close to zero. This last property aims at controlling not only the relative values of gray-level differences but also their absolute values. Then, the gray-level function ($f$) that we can consider has to:

- be decreasing;
- spread not significant gray-level differences so that uniform distribution will be acceptable.

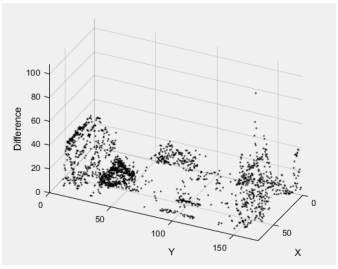
In this study, two $f$ functions, denoted here by $f_1$ and $f_2$, were evaluated: $\forall x \geq \tau, f_1(x) = \frac{1}{x-\tau}, f_2(x) = 1 + \tanh(\tau - x), \forall x < \tau, f_1(x) = f_2(x) = +\infty$. The parameter $\tau$ has been introduced to allow us to control the number of points considered in the following steps of the algorithm. Indeed, the values of $\Delta I$ which are lower than $\tau$ will result in an infinite distance (cf. Section 3.3) so that they are simply discarded; therefore, the higher the $\tau$ the less points the algorithm considers. Note that, in order to remain parameter free (but at the expense of memory and computational resources), one can set this parameter equal to zero.



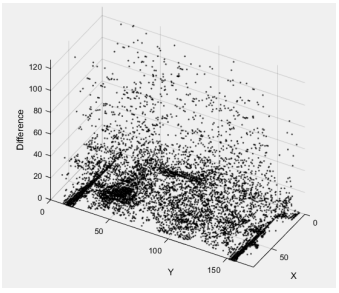

(a) (b)

Fig. 2: Comparison for $\tau = 3.0$ between (a) $f_1(x) = \frac{1}{x-\tau}$ and (b) $f_2(x) = 1 + \tanh(\tau - x)$.

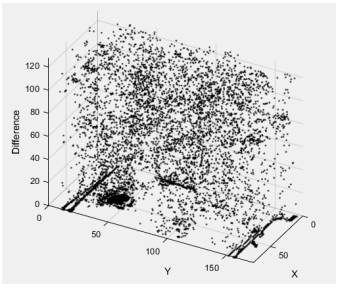
We focus on these two functions since, while having the desirable properties, the resulting value spread is quite different. The inverse function gives a gradual decline while the tanh function provides a more sudden drop for high values (Fig. 2). The point clouds provided considering respectively each of the two functions are shown in Fig. 3; both provide the expected discrepancy of the difference values, the tanh function result appearing somehow more uniformly distributed. Therefore, in




(a)




(b)




(c)

Fig. 3: 3D point cloud (a) before applying the function $f$, (b) after applying $f_1(x) = \frac{1}{x-\tau}$ and (c) after applying $f_2(x) = 1 + \tanh(\tau - x)$. The vertical axis represents the (transformed) gray-level values while the other two axis originate from the 2D image plane.

the following of the paper, we focus on tanh function that allows for better consistency with the considered naive model.

### *3.3. Distance between two points*

The cluster detection using the a-contrario approach is based on point distance [45]. In order to take into account both spatial proximity and (transformed) gray-level differences, the proposed distance is a weighted sum of two terms: the 2D spatial distance and a term representing the modified gray-level of each point. Thus, the proposed distance is defined as the minimal path length among the paths relating two points and passing through the $z = 0$ plane with $z$ being the gray-level axis. In this way, we can enforce that points with higher gray-level values after $f$-transformation be considered farther apart compared to points with lower $f$-transformed gray-level values. Since the gray-level values and spatial distance are inherently in different scales, we use the scale factor $c \in \mathbb{R}_+$ to control the weight of the spatial term with respect to the gray-level one in the distance definition. The choice of $c$ value is further discussed in Section 4.

Denoting by $y_i$ the value at pixel $i \in \mathcal{P}$, by $z_i = f(y_i)$ its transformed gray-level value and by $D_{sp}(i, j)$ the 2D spatial distance between the locations of pixels $i$ and $j$, $\forall (i, j) \in \mathcal{P}^2$, if $i = j, D(i, j) = 0$, and otherwise

$$D(i, j) = \sqrt{\Big(D_{sp}(i, j)\Big)^2 + c \times \Big({z_i}^2 + {z_j}^2\Big)}. \tag{1}$$

Let us specify that, without the constraint "if $i = j, D(i, j) = 0$", $D$ would be only a metametric since, if the symmetry property and triangle inequality hold (they can be easily demonstrated provided that $D_{sp}$ also check triangle inequality), the identity of indiscernibles does not. Indeed, $D(i, j) = 0 \Rightarrow i = j$, but the opposite ($i = j \Rightarrow D(i, j) = 0$) is not true but for pixels such that $z_i = z_j = 0$.

Finally, a cluster $C \subseteq \mathcal{P}$ is defined as a set of close points with respect to the distance value $d$: $C \subseteq \mathcal{P}$ is the set of points $i$ such that $\forall i \in C, \begin{cases} \exists j \in C \text{ s.t. } D(i, j) \leq d, \\ \forall j' \in \mathcal{P} \setminus C, D(i, j') > d. \end{cases}$

Note that for a given $d$ there may be several distinct clusters $C_i$ satisfying the previous definition. Inversely, for a given cluster $C$, there is a range of distances leading to $C$ that allows us to associate an inner border and an outer border to cluster $C$. In the following, we denote $d_{min}(C)$ and $d_{max}(C)$ the bounds of this interval.

### *3.4. Number of False Alarms*

The Number of False Alarms (NFA) is based on the considered naive model which in our case is the uniform distribution:

**Definition 1** (Naive model $\mathcal{M}$)**.** *The set of points $\mathcal{S}$ is a random set of $|\mathcal{S}|$ independent uniformly distributed variables over the 3D (2D+gray-level) space of the image.*

Note that a key point of a-contrario approaches is that the naive model does not have to be accurate, but it only has to be contradicted in the case of the target structured data (wear in our application).

The Number of False Alarms is computed by extending the NFA proposed in [7] for 2D cluster detection. Considering here a 3D space, the 2D surface areas are replaced by 3D volumes and the 2D distance by the distance defined in Eq. (1) so that, for any cluster $C$ of 3D (2D+gray-level) points,

$$NFA_{\mathcal{M}}(C, M) = N_{test} \sum_{i=k}^{M} \binom{M}{i} \underline{V}_C^i (1 - \overline{V}_C)^{M-i}, \tag{2}$$

where $k$ is the number of points in the cluster, $M$ is the total number of points and $\underline{V}_C$ and $\overline{V}_C$ are the lower and upper bounds of the relative volume of the cluster with respect to the whole image cube volume. Therefore $1 - \overline{V}_C$ represents the volume of the region that is definitely outside the cluster while $\underline{V}_C$ represents the volume of the region which is definitely inside the cluster. These volumes are obtained by relying on morphological operations as specified in the next section. Finally, $N_{test}$ that is a normalization term equivalent to the number of tests coefficient is set as a constant so that it does not impact the NFA minimization and can be discarded if one is interested only in the cluster ordering with respect to their NFA-based meaningfulness.

### *3.5. Calculating the lower and upper volumes*

Let us first define the distance between a cluster $C$ and a single point $i$ as the minimum distance between all the points in $C$ and $i$.

Then, let $\delta$ be the maximum distance between any pair of points inside the cluster $C$, and $\delta'$ be the distance between $C$ and the closest point outside $C$. The lower and upper volumes of $C$ are computed by performing a 3D mathematical morphological dilation [46]:

- The lower region is the dilation of the union of the points in $C$ by a ball structuring element having the radius $\delta/2$. Let us then denote by $\underline{V}_C$ the volume of this region divided by the volume of the image cube.

- The upper region is the dilation of the union of the points in $C$ by a ball structuring element having the radius $\delta'$. Let us then denote by $\overline{V}_C$ the volume of this region divided by the volume of the image cube.

It is worth noting that since we have used a modified distance formula (Eq. 1), all dilation operations have to be done using this custom distance. Further details will be discussed in Section 4.

### *3.6. Most meaningful clusters*

After each cluster has an assigned NFA, we compute the meaningfulness for each cluster:

$$\mathcal{S}_{\mathcal{M}}(C, M) = -\log(NFA_{\mathcal{M}}(C, M)). \tag{3}$$

In the following, only comparing cluster significance values at given value $M$ and naive model $\mathcal{M}$, we shorten significance notation as $\mathcal{S}(C) = \mathcal{S}_{\mathcal{M}}(C, M)$.

By construction of the minimum spanning tree, for any pair of considered clusters $C$ and $\mathcal{K}$, either $C \cap \mathcal{K} = \emptyset$ or $C \subsetneq \mathcal{K}$ or $\mathcal{K} \subsetneq C$. Then, to avoid redundant results (detection of the same cluster several times), we focus on *maximal* clusters such that a cluster $C \subset \mathcal{P}$ is said *maximal* if $\forall \mathcal{K} \subsetneq C, \mathcal{S}(\mathcal{K}) < \mathcal{S}(C)$ and $\forall \mathcal{K} \supsetneq C, \mathcal{S}(\mathcal{K}) \leq \mathcal{S}(C)$ [7].

In Algorithm 1, meaningfulness *maximality* is handled as a constraint: to be added to the list of *maximal meaningful* clusters $\mathbf{C}$, a cluster must not intersect any of the clusters already in $\mathbf{C}$.

## 4. Implementation

The implementation process (Fig. 1) begins with an application dependent step. This step can be modified to adapt our proposed algorithm to new applications. Once the acquisition is complete, it is imperative to make sure the two images we compare are spatially co-registered and have the same general illumination level. Indeed, the images can be taken in different conditions, for which the most relevant influencing factors are the lighting and/or the field of view. In Fig. 1, these preprocessing steps are grouped in a block called "Acquisition dependent steps".

The next block deals with the computation of $\Delta I$Since the UVIFL images are classic RGB images (cf. Section 5.2), we have employed the CIEDE2000 formula, after projecting the RGB values into the CIELAB color space. Alternatives may be considered to create the difference map, from a simple Euclidean distance between RGB values to more sophisticated color difference models. However, in this study, the employed difference formula provides consistent results which agree with the expected wear areas.

The third block is the proposed change detection algorithm that applies to any difference image provided by block 2 and used for finding the clusters and computing their significance. The process is outlined step by step in Algorithm 1, as follows.

First, we create a minimum spanning tree of the points derived from the difference image based on the distances computed using Eq. (1). The spanning tree is constructed as follows [7]: we initialize a graph whose nodes are finite-coordinate points, there are no edges, and all distances between pairs of points are pre-computed. The spanning tree is then constructed in an iterative way, during which, at each iteration (i) we select the two nearest nodes among the unconnected nodes and (ii) we create an edge between these two nodes. For an easier derivation of clusters of connected points, we also introduce a hierarchical representation of this iterative process, in which, at each iteration, the nearest nodes are merged in a parent node which stores as well the maximum distance between pairs of points belonging to the node. That is to say, for any created parent node, this additional stored value is simply the largest distance between any of the children nodes. In the algorithm, we call a subtree a node along with all its children nodes.

Each parent node is then considered as a potential cluster. As mentioned earlier, for each cluster, we compute two separate dilations, one being performed for the lower bound and one for the upper bound region. These dilations are done in 3D by using Eq. (1) as the distance. As expected, this step produces different shapes than those obtained by using the standard dilation based on a 3D-ball structuring element or Euclidean distance. The volumes of these regions are involved in the computation of the NFA and of the meaningfulness values for each cluster. For saving computational time (and since preliminary tests did not show a difference in the derived ordering of the clusters), we only compute NFA up to scale owing to $N_{test}$ which boils down to deriving meaningfulness values up to a shift. The final step is to find the maximal clusters. Starting with an empty set $\mathbf{C}$ and the list of clusters $C_i$ ranked in decreasing order of meaningfulness, at each step, we increment $i$ to select the next cluster $C_i$, and, only if it is disjoint from any cluster already stored in $\mathbf{C}$, we add it to $\mathbf{C}$. In the end, $\mathbf{C}$ represents the output of the algorithm which is the ranked list of detected clusters based on their $\mathcal{S}$ value.

Let us finally provide some practical notes:

- The choice of the parameter $c$ in Eq. (1) is driven by both the data features and the expected clusters. Indeed, when $c \gtrsim 0$, the clusters would be spatially large and very dense including pixels almost whatever their gray-level transformation value. Conversely, a value $c \gg 0$ favors clusters spatially scattered and very sparse only including pixels with very low values (after transformation). Then, $c$ parameter is set based on image resolution spatial and radiometric features. Based on performed experiences, we set $c = 0.1$ as default value (used in all our experiments).

- When calculating Eq. (2), intermediate values of $\binom{M}{k}$ can get very large and generate overflows. In cases with a small $M$ (around 2000 or lower), we can deal with this by using Big number data types. Otherwise, instead we can approximate Eq. (3) using the Hoeffding approximation like in [47].

**Algorithm 1** Change detection between the current frame and the reference frame.

Perform the pre-processing for $I_0$ and $I$
Compute the color difference map $\Delta I$ between $I$ and $I_0$
**for** each pixel $j$ in $\Delta I$ **do**
    $\Delta I(j) = f(\Delta I(j))$
**end for**
$\mathcal{P} \leftarrow$ 3D points derived from pixels $j$ such that $\Delta I(j) < +\infty$
$M \leftarrow |\mathcal{P}|$
**for** each pair of points $i$ and $j$ in $\mathcal{P}^2$ **do**
    Compute $D(i, j)$ according to Eq. (1)
**end for**
Compute the minimum spanning tree for the $\mathcal{P}$ points based on $D(i, j)$ so that each subtree in hierarchical representation stores in its root the maximum distance between its points
$V_{\mathcal{P}} \leftarrow$ volume of image cube
**for** each subtree $T$ **do**
    $C \leftarrow$ cluster of points in $T$
    $\delta \leftarrow$ value stored in the root of $T$
    $\delta' \leftarrow$ value in the parent node of $T$
    $\underline{V}_C \leftarrow [volume\ of\ dilate(C, \delta/2)]/V_{\mathcal{P}}$
    $\overline{V}_C \leftarrow [volume\ of\ dilate(C, \delta')]/V_{\mathcal{P}}$
    $k \leftarrow$ the number of points in $C$
    Compute NFA value (up to scale owing to $N_{test}$) according to Eq. (2) using values $k$, $M$, $\underline{V}_C$ and $\overline{V}_C$
    $\mathcal{S} \leftarrow -\log(NFA)$
**end for**
$\mathcal{J} \leftarrow$ list of indices of the clusters sorted according to $\mathcal{S}$
$\mathbf{C} \leftarrow \emptyset$
**for** each index $j$ in $\mathcal{J}$ **do**
    $C_j \leftarrow j^{th}$ cluster according to $\mathcal{J}$
    **if** $\forall C_l \in \mathbf{C}, C_j \cap C_l = \emptyset$ **then**
        $\mathbf{C} \leftarrow \mathbf{C} \cup \{C_j\}$
    **end if**
**end for**
$\mathbf{C}$ is the list of detected clusters

- Decreasing the quantization level for the gray-level values (e.g 128 levels instead of 256) can help improving the performance by reducing the number of calculations in the 3D morphological operations.

- In practice, the parameter $\tau$ can also be set to higher values to reduce the number of points $M$ (cf. Eq. (2)). This will not affect the output as long as the omitted points have color difference values less than the minimum amount of difference perceivable (this value depends on the application, amount of noise present and the color difference formula used). In all our experiments, we set $\tau = 3$ which allows us to focus on only 20% of the pixels (which is still much more important that the wear areas that represent only up to a few percents of the whole image).

## 5. Results

### *5.1. Evaluation of the algorithm robustness on simulated data*

In this section, we use a simulated difference image to evaluate the performance of the proposed algorithm. First, we created a binary map which serves as ground truth to distinguish the background and the foreground. The foreground region was hand drawn to represent two clusters with complex shapes. Then, the pixel values in both regions were randomly drawn assuming two different heavy-tailed distributions. We focus on the case of a heavy-tailed distribution since a) it simulates the high value noise present in the background that is more disruptive than Gaussian noise for instance and b) it allows us to illustrate that the naive model $\mathcal{M}$ (Def. 1) does not need to be exact. Specifically, since the drawn values have to simulate color difference values that are strictly non-negative, we focused on a Nakagami distribution base in both cases (however any similar distributions with the adequate parameters can be used here). Specifically, for the foreground, we use a Nakagami distribution plus a *shift* value which allows us to easily control the *mean* without changing the shape of the distribution. To represent realistic data with respect to our application, the background is a spread out distribution near zero and a tail with high values while the foreground, much less spread out, has a higher mean value than the background. Then, the aim is to detect the two clusters present in the image and to evaluate the result using the binary map as the ground truth.

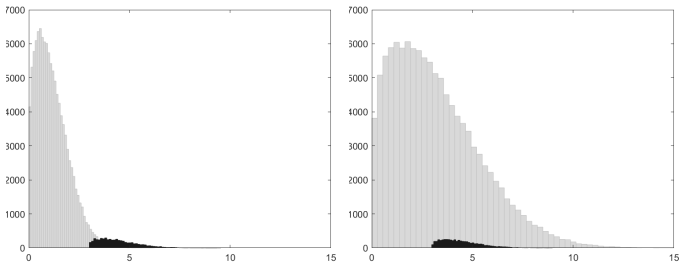
For a quantitative evaluation of the proposed algorithm, we simulate data considering a large range of mean and standard deviation values for the foreground and the background. Specifically, we consider the two following experiments:

- Experiment 1: we increase progressively the spread of the background with respect to fixed foreground distribution parameters;

- Experiment 2: we vary the mean value of the foreground with respect to fixed background distribution parameters.

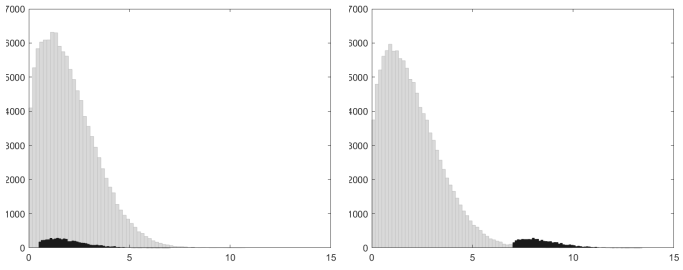
The Nakagami distribution has two parameters, denoted $\mu$ and $\omega$, which control the *shape* and *spread* of the distribution, respectively. In the first experiment, we change the spread $\omega_0$ of the background (from 2.0 to 11.0 by steps of 0.5; and then to 15 by steps of 1.0) while keeping the shape of the background $\mu_0$ constant. The parameters of the foreground ($\omega_1$ and $\mu_1$) are also kept constant (Fig. 4). From the *shape* and *spread* of the distribution, we can derive the mean of the background that is found to vary from 1.16 in the first step to 3.19 in the last. Since the mean of the foreground is constant at 4.42 we expect the detection to become progressively harder as the two means get closer.

In the second experiment, given constant parameters for both regions, the distribution of the foreground is progressively shifted towards higher values, which results in a gradual increase of the mean of the foreground from near zero to higher values. This allows us to evaluate the performance with respect to the overlapping between background and foreground distributions. Specifically, the background shape $\mu_0$ is set to 0.6 and $\omega_0$ to 6.0. This results in a mean of nearly 2.0 and a standard



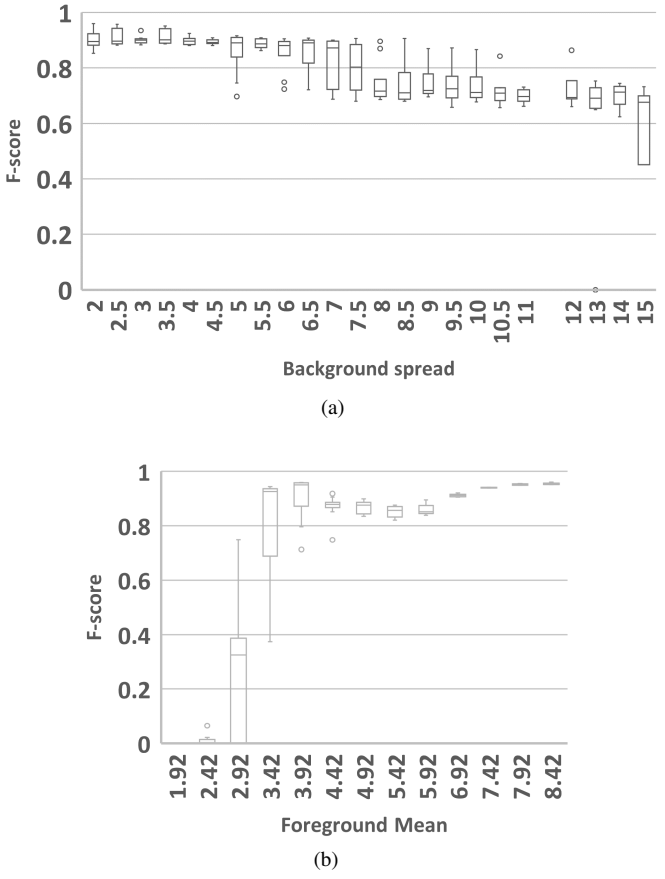
(a) (b)

Fig. 4: Experiment 1: a) histogram of the foreground (gray) and background (black) in the first step and b) the last step.



(a) (b)

Fig. 5: Experiment 2: a) histogram of the foreground (black) and background (gray) in the first step and b) the last step.




(a)

(b)

Fig. 6: The F-score for the results of the algorithm with different (a) spread for the background (experiment 1) and (b) mean for the foreground (experiment 2).

deviation of nearly 1.4 . For the foreground, the mean will progressively increase from 1.92 to 8.42 in steps of 0.5 (Fig. 5).

In both experiments, for each simulated image corresponding to a gradual change of parameters, we apply our algorithm to detect the clusters. Besides, to get statistically significant results, 10 image realisations are considered for each given set of distribution parameters. From the ground truth binary map, the detection results are evaluated in terms of F-score that is computed from *precision* and *recall* values as follows: $F_1 = \frac{precision \times recall}{precision + recall}$.

The evaluation results using the F-score are shown in Fig. 6 for each experiment. In each step, the range observed during the repetitions, along with their average and outliers (if any) have been shown. According to these charts, experiment 1 produces an F-score (on average) higher than 0.8 until step 12 for which the mean of the background is 2.24; and higher than 0.7 until step 20 for which the mean of the background is 2.85. In experiment 2, the algorithm provides F-scores (in average) higher than 0.8 from step 4 for which the mean of the foreground is 3.42. Considering the fact that the background is chosen to not represent the naive model, thus providing a greater challenge, the F-score plummets only in cases with extreme amount of high value pixels in the background.

Figures 7 and 8 show examples of the output of the algorithm for each experiment. Ideally, in each frame two separate clusters should be detected, namely one bigger to the left and one smaller. In addition, we expect that the red colour, which indicates the most significant cluster, highlight the bigger one. In both cases, the algorithm shows resilience to the presence of the noise until the background and foreground become indistinguishable from each other.

These two experiments show the resilience of the algorithm to background noise and how well it can detect minute differences between the background and foreground. It is worth mentioning that these simulations aim at evaluating the proposed approach in worst scenario cases, since we expect actual data be less noisy and/or pre-processed by a noise removal process. In the next section, we continue the evaluation by using real images.

### *5.2. Performance on actual data*

#### *5.2.1. Dataset description*

In this study, we used multi-temporal UV induced fluorescence (UVF or UVIFL) images collected in the "Violins UV-IFL imagery" dataset[1] [48]. UVIFL photography is a non-invasive diagnostic technique, widely adopted in Cultural Heritage studies, that allows us to see details not perceivable using visible light [49]. It exploits the characteristic of some substances present on the varnishes that absorb UV radiations re-emitting light in the visible spectrum. In the case of historical violins, UVIFL images are generally used to highlight possible

[1] https://vision.unipv.it/research/UVIFL-Dataset/

(a) (b) (c) (d) (e) (f)

Fig. 7: Experiment 1: the detected clusters from (a) to (f) in steps 1,8,12,17,21 and 27. (a) shows the perfect segmentation.

(a) (b) (c) (d) (e) (f)

Fig. 8: Experiment 2: the detected clusters from (a) to (f) in steps 1,2,3,6,9 and 14. (f) shows the perfect segmentation.

restorations or interesting areas [50] where to apply more precise but slower diagnostic techniques, like XRF [51] or FTIR spectroscopy [52]. The capability to see "hidden" details of a surface is crucial in our scenario, since it increases the chances of an early detection of new alterations. In fact, visible light can be deceptive. When a new alteration is clearly identifiable under a standard illumination, it is likely that an irreversible damage to the surface varnish has already occurred. Moreover, substances used for cleaning the instruments can temporarily hide the presence of alterations, and, thus, slow down the detection. On the contrary, the effectiveness of UVIFL photography to characterize variations in the varnishes is well known in the Cultural Heritage field [53].

The "Violins UVIFL imagery" dataset contains UVIFL images of both historical and sample violins. Regular acquisitions have been performed on two historical violins held in Museo del Violino in Cremona (Italy), "Carlo IX" (c.1566) made by Andrea Amati and "Vesuvio" (1727) made by Antonio Stradivari. However, as good news concerning them, they did not show any new wear areas (only "Vesuvio" showed a very slight alteration on its back plate). Thus, for wear monitoring purpose, we considered three artificially created sample sequences containing images of artificially altered samples for the study of various possible alterations over a long-term use.

The alterations were created scrubbing the surface with a cloth damped with alcohol to reproduce, as faithfully as possible, the effect of mechanical wear during playing. The alteration process was repeated multiple times. At each step we took (at least) three photos of the samples, for safety, to exclude errors due to accidental wrong acquisitions. During the experiments we considered only the best shot for each time frame.

The first artificial sequence, called WS01 (Fig. 9(a) and (b)), is a wood sample which simulates an alteration in an area with intact varnish. This set contains one reference image of the initial state of the sample and 20 altered frames.

The second artificial sequence, called WS02 (Fig. 9(c) and (d)), is a wood sample which simulates an alteration in an area with a thin layer of varnish. This set contains one reference image of the initial state of the sample and 8 altered frames.

The third sequence, called SV01 (Fig. 9(e) and (f)), contains images of the lower part of the back plate of a sample violin. This set simulates the growing of wear starting from an area already ruined and consists of one reference image and 20 altered frames.

All the images were acquired following a rigorous acquisition protocol designed to minimize, as much as possible, the presence of ambient noise [54, 55]. The wood samples were placed on a small support to maintain them stable during the shot, while the instruments were placed on an ad-hoc rotating platform that allow us to move them precisely at the needed angle. The photos were taken with a Nikon D4 full-frame digital camera with a 50 mm f/1.4 Nikkor objective, 30s exposure time, aperture f/8, ISO 400. We used two wood lamp tubes (Philips TL-D 36 W BBL IPP low-pressure Hg tubes, 40 Watt, emission peak $\sim 365nm$) as UV-A lighting source. The lamps were oriented at 45 degree to uniformly illuminate all the surface of the samples/instruments. Note that, even with such a rigorous acquisition protocol, some noise can still occur, especially in the most rounded part of the violins.

The image preprocessing includes spatial registration performed as follows. We rely on extracting and matching SIFT [56] features in the original image and in the subsequent frames. These matched points are then used to estimate a specific transformation between a given couple of frames using the robust estimation [57]. Since our captures have been performed on a rigid object in a controlled environment, this on-shelf approach proved to provide alignment at pixel level that is sufficient for the following process given the expected scale change.

Finally, to make the evaluation of the algorithm possible, the ground truth for each frame has been created manually from careful visual inspection. In the absence of any exact (due intrinsically to the wear construction process) knowledge on the boundaries of wear regions, they have been defined intentionally very loose.

### *5.2.2. Evaluation regarding 3D clustering*

Figure 10 shows the result of the proposed NFA-based clustering for four sample frames of WS01. In each frame the top 4 significant clusters have been indicated by their borders. The blue border shows the most meaningful cluster. As we can see, small noises change from frame to frame, big artefacts have a constant size and location, and the wear area grows over time.

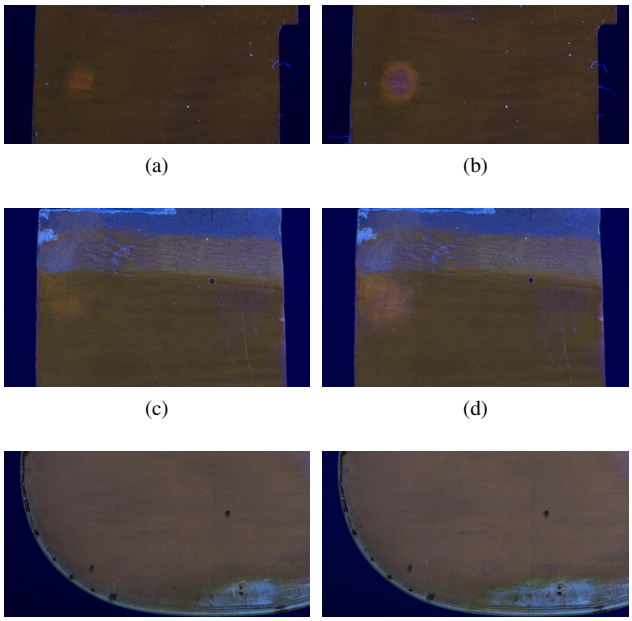


Fig. 9: Samples of UVIFL images contained in the dataset (for clarity we have chosen those where the alterations are particularly visible): (a) and (b) from set WS01, (c) and (d) from set WS02, (e) and (f) from set SV01.

In all cases, small noises have been ignored and significant high change areas have been identified.

We have compared the clustering output of our algorithm with several other clustering methods, namely: Agglomerative hierarchical clustering with complete linkage [58], Kmeans++ [59], robust spectral clustering [60], Expectation Maximization for Gaussian Mixture models (EM GM) [61], clustering by fast search [62], GBKmeans [63], clustering by local gravitation [64] and HDBSCAN [65]. Thus, the chosen algorithms belong to different families of clustering methods and include both established approaches and more recent works. Each of these algorithms have their own strength and weakness features such as their robustness to noise and their need to get the number of clusters as input. To make the comparison feasible we have set the number of clusters $k$ equal to 4 whenever needed. Also, we have chosen manually the best cluster in the output to ensure to get an upper bound for performance metrics of that algorithm.

Considering every image sample present in the three considered sequences (Seq. WS01, WS02, SV01), the precision, recall and F-score metrics have been computed for each clustering result obtained on the point cloud $\mathcal{P}$ generated considering the pixels with $\Delta I$ values greater than $\tau$ (for fair comparison with our approach). Table 1 summarizes the obtained results in terms of the average and standard deviation of F-score values for each sequence. The NFA clustering performs well for all three sequences (high average F-score); and maintains that performance for each frame in the sequence (very low standard deviation).

Furthermore, it appears that the performance of each clustering method (except the proposed NFA-based one) varies from one sequence dataset to another and the best alternative to our algorithm is different for each sequence. This is due to the volatile nature of the noise and artifacts present in our data. Therefore, it seems difficult to choose, among previous works, a performing clustering method that manages successfully and consistently the different kinds of noise.

Table 1: Average and standard deviation of F-score values for different clustering algorithms on Seq. WS01, WS02 and SV01. Best results are in bold, second best results are underlined.

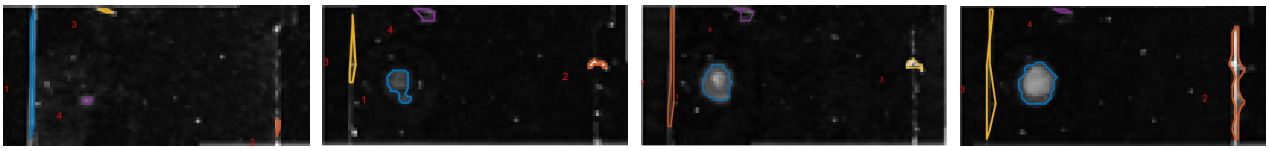
| | WS01 | | WS02 | | SV01 | |
|---|---|---|---|---|---|---|
| **Algorithm** | **Avg** | **Std** | **Avg** | **Std** | **Avg** | **Std** |
| **Agglomerative [complete] [58]** | 0.5318 | 0.0935 | 0.6160 | 0.0848 | 0.6065 | 0.1023 |
| **Kmeans++ [59]** | 0.4914 | 0.1129 | <u>0.6769</u> | 0.0665 | 0.5930 | 0.0912 |
| **Robust Spectral [60]** | 0.6916 | 0.1479 | 0.6475 | 0.0665 | 0.6753 | 0.1177 |
| **EM GM [61]** | 0.5668 | 0.1437 | 0.6379 | 0.0798 | <u>0.7503</u> | 0.1157 |
| **Fast Search [62]** | 0.4227 | 0.1150 | 0.5273 | 0.0620 | 0.5828 | 0.0799 |
| **GBKmeans [63]** | 0.3715 | 0.0858 | 0.4870 | 0.0559 | 0.5760 | 0.1051 |
| **Local Gravitation [64]** | <u>0.6939</u> | 0.0902 | 0.6569 | 0.1213 | 0.6139 | 0.0936 |
| **HDBSCAN [65]** | 0.6272 | 0.1319 | 0.5968 | 0.0477 | 0.6418 | 0.0866 |
| **NFA(Alg. 1)** | **0.8493** | 0.0492 | **0.7634** | 0.0382 | **0.8018** | 0.0530 |

### *5.2.3. Comparison with 2D segmentation*

An alternative to clustering is $\Delta I$ image segmentation (to detect the altered areas) followed by labelling of cluster components. Therefore, we also evaluate our proposal against a two-step method: a binary segmentation, namely FRFCM [66], followed by a $2D$ clustering, namely HDBSCAN [65].

FRFCM is a modified fuzzy c-means algorithm that incorporates local spatial information by using morphological reconstruction, which improves the classic fuzzy c-means to help in dealing with the different types of noise. Applied to our data, it provides rather good separation between background and foreground. Then, to spatially cluster the points produced from FRFCM we use HDBSCAN (Hierarchical Density-Based Spatial Clustering with Application with Noise) which is a framework for density-based cluster analysis. The algorithm provides a complete clustering hierarchy of all possible density based clusters and - crucially for us - it also provides a global, optimal non-hierarchical solution which maximizes the overall stability of the proposed clusters. This means we can evaluate our automatic process (regarding the number of clusters) with a direct comparison of resulting clusters from both methods. In addition, for each frame, the closest cluster to the ground truth is selected manually for the calculation of the performance metrics.

Tables 2, 3 and 4 show three sample frames from each dataset: their ground truth and output of each method. Also, for comparison with earlier attempts at wear detection on the same dataset, we included the results obtained from the process proposed in [48]. This solution was based on histogram quanti-



(a) (b) (c) (d)

Fig. 10: Clustering output from frames 3, 9, 15 and 20 of set WS01 using the proposed NFA clustering (Algorithm 1).

Table 2: Comparison between the proposed NFA clustering, Dondi et al. [48], FRFCM+HDBSCAN clustering and the ground truth for some sample frames from set WS01.

| No. | Ground truth | NFA clustering (Ours) | Dondi et al. [48] | FRFCM+HDBSCAN |
|---|---|---|---|---|
| $S_1:9$ |  |  |  |  |
| $S_1:15$ |  |  |  |  |
| $S_1:20$ |  |  |  |  |

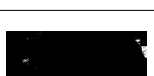
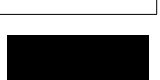
Table 3: Comparison between the proposed NFA clustering, Dondi et al. [48], FRFCM+HDBSCAN clustering and the ground truth for some sample frames from set WS02.

| No. | Ground truth | NFA clustering (Ours) | Dondi et al. [48] | FRFCM+HDBSCAN |
|---|---|---|---|---|
| $S_3:4$ |  |  |  |  |
| $S_3:7$ |  |  |  |  |
| $S_3:9$ |  |  |  |  |

zation and genetic algorithm, and was designed to minimize the false positive detection and to quickly give a rough estimation of the likely position of the altered region(s).

A qualitative analysis of the results from all three methods indicates that we have successfully dealt with background noise and artifacts from UV reflections in the majority of cases. For example, in the sequence SV01, reflections on the border of the violin are very close to the actual wear region. The NFA clustering has managed to avoid them (almost) completely, while FRFCM+HDBSCAN have grouped them together with the wear in a few cases. We have also improved the results with respect to [48], that, even if generally less prone to false detection than FRFCM+HDBSCAN, is also less effective that NFA in properly identifying the boundaries of the altered regions. Inherently the NFA clustering allows for controlling the number of false alarms. This results in globally better wear detection (lower number of false positives).

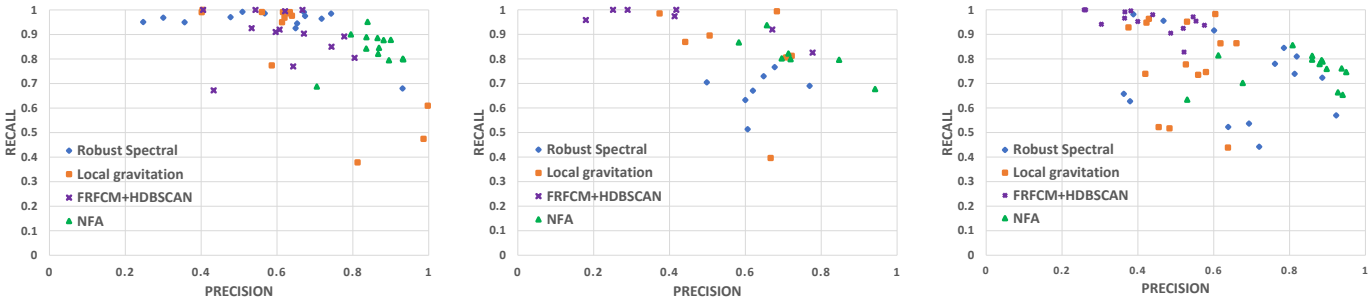
Table 4: Comparison between the proposed NFA clustering, Dondi et al. [48], FRFCM+HDBSCAN clustering and the ground truth for some sample frames from set SV01.

| No. | Ground truth | NFA clustering (Ours) | Dondi et al. [48] | FRFCM+HDBSCAN |
|---|---|---|---|---|
| $S_2:9$ |  |  |  |  |
| $S_2:15$ |  |  |  |  |
| $S_2:20$ |  |  |  |  |

To summarize both experiments, we compare the performance of two of the 3D clustering methods mentioned above (Local Gravitation [64] and robust spectral clustering [60]); FRFCM+HDBSCAN [66, 65] and our proposal (Alg. 1). Figure 11 shows the precision/recall charts for sequences WS01, WS02 and SV01. In all three sets, the proposed NFA clustering has better precision while maintaining an acceptable recall in most cases. As we can see, the FRFCM+HDBSCAN method tends to have good recall values but with poor precision i.e high false positives. This is due to the fact that the binary segmentation step only filters out the low value noise present in the image. Therefore, in the clustering step, the high value artifacts are hard to separate from the actual wear. As a result, it is vital to consider the gray-level values at the same time as the spatial information.

### *5.2.4. Computational cost*

We use a MATLAB vanilla implementation of the algorithm, and the cost is divided evenly between morphological operations and the computation of the NFA value. In the current non-optimized form, the computational time is roughly equal to the one of robust spectral clustering, which is however optimized. For a more time-constraint application, our algorithm could thus benefit straightforwardly from parallelization and a better implementation of the morphological operations (e.g. by relying on run-length encoding [67]).




(a) (b) (c)

Fig. 11: Precision-Recall plot for WS01 (a), WS02 (b) and SV01 (c). For a given algorithm (indicated by the color), each point highlights the performance at a specific time-step of the sequence.

## 6. Conclusion

In this work, we proposed a probabilistic algorithm to detect clusters of change between two temporally different images of the same scene. Our proposal is based on an a-contrario framework and performs the clustering process directly on the grey-level difference image, while dealing with the background noise and artifacts. Series of test cases generated by random numbers, used to test the limitations of the method, show flexibility to background noise and the ability to detect minute differences. Also, comparisons with recent clustering methods show meaningful improvements while having the benefit of an inherent ranking criterion for the resulting clusters.

This approach can be used in preventive conservation as a fast, preliminary examination of the surface of a violin able to identify the most likely altered areas. Thus, a verification using more precise but slower techniques (like spectroscopic analyses) will be done only on the detected areas, reducing the time needed for completing the monitoring procedures.

For future studies, we intend to perform a long-term monitoring process considering real historical violins which are played weekly. Beside creating a valuable dataset for our community, this will allow us to validate our algorithm as well on real wear patterns. In addition, the current version of the algorithm can benefit from parallel implementation and perhaps alternative methods for faster 3D dilation operations. Lastly, while in the current work we have only used one pair of images at a time to track the changes, it is interesting and useful to incorporate time information in our model. Inherently, the evolution of the wear region through time is different from that of artefacts being present in the image; this can help us to differentiate wear regions better. In addition, in an ongoing monitoring process, it can help us detect a newly created wear at an earlier stage.

## References

[1] S. Bradley, Preventive conservation research and practice at the british museum, Journal of the American Institute for Conservation 44 (2005) 159–173.

[2] E. Lucchi, Review of preventive conservation in museum buildings, Journal of Cultural Heritage 29 (2018) 180 – 193.

[3] G. V. Fichera, M. Albano, G. Fiocco, C. Invernizzi, M. Licchelli, M. Malagodi, T. Rovetta, Innovative monitoring plan for the preventive conservation of historical musical instruments, Studies in Conservation 63 (2018) 351–354.

[4] T. Rovetta, C. Invernizzi, G. Fiocco, M. Albano, M. Licchelli, M. Gulmini, G. Alf, D. Fabbri, A. Rombolà, M. Malagodi, The case of antonio stradivari 1718 ex-san lorenzo violin: History, restorations and conservation perspectives, Journal of Archaeological Science: Reports 23 (2019) 443–450.

[5] P. Dondi, L. Lombardi, M. Malagodi, M. Licchelli, Automatic identification of varnish wear on historical instruments: The case of antonio stradivari violins, J. Cult. Heritage 22 (2016) 968–973.

[6] A. Desolneux, L. Moisan, J.-M. Morel, Meaningful alignments, International journal of computer vision 40 (2000) 7–23.

[7] A. Desolneux, L. Moisan, J.-M. Morel, A grouping principle and four applications, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (2003) 508–513.

[8] A. Rezaei, S. Le Hégarat-Mascle, E. Aldea, P. Dondi, M. Malagodi, One step clustering based on a-contrario framework for detection of alterations in historical violins, in: 25th International Conference on Pattern Recognition (ICPR2020), 2021, p. 9348 – 9355.

[9] A. Varghese, J. Gubbi, A. Ramaswamy, P. Balamuralidhar, Changenet: A deep learning architecture for visual change detection, in: Proceedings of the European Conference on Computer Vision (ECCV) Workshops, 2018, pp. 0–0.

[10] Q. Wang, Z. Yuan, Q. Du, X. Li, Getnet: A general end-to-end 2-d cnn framework for hyperspectral image change detection, IEEE Transactions on Geoscience and Remote Sensing 57 (2018) 3–13.

[11] L. Khelifi, M. Mignotte, Deep learning for change detection in remote sensing images: Comprehensive review and meta-analysis, IEEE Access 8 (2020) 126385–126400.

[12] M. Mandal, S. K. Vipparthi, An empirical review of deep learning frameworks for change detection: Model design, experimental frameworks, challenges and research needs, IEEE Transactions on Intelligent Transportation Systems (2021).

[13] M. D. Li, K. Chang, B. Bearce, C. Y. Chang, A. J. Huang, J. P. Campbell, J. M. Brown, P. Singh, K. V. Hoebel, D. Erdoğmuş, et al., Siamese neural networks for continuous disease severity evaluation and change detection in medical imaging, NPJ digital medicine 3 (2020) 1–9.

[14] M. Sturari, M. Paolanti, E. Frontoni, A. Mancini, P. Zingaretti, Robotic platform for deep change detection for rail safety and security, in: 2017 European Conference on Mobile Robots (ECMR), IEEE, 2017, pp. 1–6.

[15] S. Verma, A. Panigrahi, S. Gupta, Qfabric: Multi-task change detection dataset, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 1052–1061.

[16] C. Hu, Y. Wang, An efficient convolutional neural network model based on object-level attention mechanism for casting defect detection on radiography images, IEEE Transactions on Industrial Electronics 67 (2020) 10922–10930.

[17] 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.

[18] D. Comaniciu, P. Meer, Mean shift: A robust approach toward feature space analysis, IEEE Transactions on Pattern Analysis & Machine Intelligence 24 (2002) 603–619.

[19] A. Y. Ng, M. I. Jordan, Y. Weiss, On spectral clustering: Analysis and an algorithm, in: Advances in neural information processing systems, 2002, pp. 849–856.

[20] S. Guha, R. Rastogi, K. Shim, Rock: A robust clustering algorithm for categorical attributes, Information systems 25 (2000) 345–366.

[21] B. Grosjean, L. Moisan, A-contrario detectability of spots in textured backgrounds, Journal of Mathematical Imaging and Vision 33 (2009) 313.

[22] T. Veit, F. Cao, P. Bouthemy, An a contrario decision framework for region-based motion detection, International journal of computer vision 68 (2006) 163–178.

[23] T. Veit, F. Cao, P. Bouthemy, Space-time a contrario clustering for detecting coherent motions, in: Proceedings 2007 IEEE International Conference on Robotics and Automation, IEEE, 2007, pp. 33–39.

[24] N. Widynski, M. Mignotte, A contrario edge detection with edgelets, in: 2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), IEEE, 2011, pp. 421–426.

[25] C. Akinlar, C. Topal, Edlines: A real-time line segment detector with a false detection control, Pattern Recognition Letters 32 (2011) 1633 – 1642.

[26] E. Aldea, S. Le Hégarat-Mascle, Robust crack detection for unmanned aerial vehicles inspection in an a-contrario decision framework, Journal of Electronic Imaging 24 (2015) 061119–061119.

[27] G. Liu, Y. Gousseau, F. Tupin, A contrario comparison of local descriptors for change detection in very high spatial resolution satellite images of urban areas, IEEE Transactions on Geoscience and Remote Sensing (2019).

[28] P. Moulon, P. Monasse, R. Marlet, Adaptive structure from motion with a contrario model estimation, in: Asian Conference on Computer Vision, Springer, 2012, pp. 257–270.

[29] E. Michaelsen, Self-organizing maps and gestalt organization as components of an advanced system for remotely sensed data: An example with thermal hyper-spectra, Pattern Recognition Letters 83 (2016) 169 – 177. Advances in Pattern Recognition in Remote Sensing.

[30] Y. Yan, J. Ren, G. Sun, H. Zhao, J. Han, X. Li, S. Marshall, J. Zhan, Unsupervised image saliency detection with gestalt-laws guided optimization and visual attention based refinement, Pattern Recognition 79 (2018) 65 – 78.

[31] J. L. Lisani, J.-M. Morel, Detection of major changes in satellite images, in: Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429), volume 1, IEEE, 2003, pp. I–941.

[32] F. Rousseau, S. Faisan, F. Heitz, J.-P. Armspach, Y. Chevalier, F. Blanc, J. de Seze, L. Rumbach, An a contrario approach for change detection in 3d multimodal images: application to multiple sclerosis in mri, in: 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2007, pp. 2069–2072.

[33] A. Robin, L. Moisan, S. Le Hégarat-Mascle, An a-contrario approach for subpixel change detection in satellite imagery, IEEE TPAMI 32 (2010) 1977–1993.

[34] A. Flenner, G. Hewer, A helmholtz principle approach to parameter free change detection and coherent motion using exchangeable random variables, SIAM Journal on Imaging Sciences 4 (2011) 243–276.

[35] C. Akinlar, C. Topal, Edcircles: A real-time circle detector with a false detection control, Pattern Recognition 46 (2013) 725–740.

[36] V. Pătrăucean, P. Gurdjos, R. G. von Gioi, Joint a contrario ellipse and line detection, IEEE TPAMI 39 (2017) 788–802.

[37] O. Martorell, A. Buades, J. L. Lisani, Multiscale detection of circles, ellipses and line segments, robust to noise and blur, IEEE Access 9 (2021) 25554–25578.

[38] G. Palma, I. Bloch, S. Muller, Detection of masses and architectural distortions in digital breast tomosynthesis images using fuzzy and a contrario approaches, Pattern Recognition 47 (2014) 2467–2480.

[39] S. Zair, S. Le Hégarat-Mascle, E. Seignez, A-contrario modeling for robust localization using raw GNSS data, IEEE Transactions on Intelligent Transportation Systems 17 (2016) 1354–1367.

[40] A. Myaskouvskey, Y. Gousseau, M. Lindenbaum, Beyond independence: An extension of the a contrario decision procedure, International journal of computer vision 101 (2013) 22–44.

[41] R. G. Von Gioi, J. Jakubowicz, J.-M. Morel, G. Randall, Lsd: A fast line segment detector with a false detection control, IEEE transactions on pattern analysis and machine intelligence 32 (2008) 722–732.

[42] S. Le Hégarat-Mascle, E. Aldea, J. Vandoni, Efficient evaluation of the number of false alarm criterion, EURASIP Journal on Image and Video Processing 2019 (2019) 35.

[43] A. Rezaei, E. Aldea, P. Dondi, M. Malagodi, S. Le Hégarat-Mascle, Detecting alterations in historical violins with optical monitoring, in: Proceedings of the 14th International Conference on Quality Control by Artificial Vision (QCAV), volume 11172, 2019, pp. 1117210–1 – 1117210–8. doi:10.1117/12.2521702.

[44] M. R. Luo, G. Cui, B. Rigg, The development of the cie 2000 colour-difference formula: Ciede2000, Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur 26 (2001) 340–350.

[45] A. Desolneux, L. Moisan, J.-M. Morel, From gestalt theory to image analysis: a probabilistic approach, volume 34, Springer Science & Business Media, 2007.

[46] P. Soille, Morphological image analysis: principles and applications, Springer Science & Business Media, 2013.

[47] F. Dibos, G. Koepfler, S. Pelletier, Adapted windows detection of moving objects in video scenes, SIAM Journal on Imaging Sciences 2 (2009) 1–19.

[48] P. Dondi, L. Lombardi, M. Malagodi, M. Licchelli, Segmentation of multi-temporal uv-induced fluorescence images of historical violins, in: New Trends in Image Analysis and Processing – ICIAP 2019, volume 11808 of *Lecture Notes in Computer Science*, Springer International Publishing, 2019, pp. 81–91. doi:10.1007/978-3-030-30754-7_9.

[49] K. Janssens, R. Van Grieken, Non-destructive micro analysis of cultural heritage materials, volume 42, Elsevier, 2004.

[50] P. Dondi, L. Lombardi, C. Invernizzi, T. Rovetta, M. Malagodi, M. Licchelli, Automatic analysis of uv-induced fluorescence imagery of historical violins, J. Comput. Cult. Herit. 10 (2017) 12:1–12:13.

[51] T. Rovetta, C. Invernizzi, M. Licchelli, F. Cacciatori, M. Malagodi, The elemental composition of stradivari's musical instruments: new results through non-invasive edxrf analysis, X-Ray Spectrometry 47 (2018) 159–170.

[52] C. Invernizzi, G. V. Fichera, M. Licchelli, M. Malagodi, A non-invasive stratigraphic study by reflection ft-ir spectroscopy and uv-induced fluorescence technique: The case of historical violins, Microchemical Journal 138 (2018) 273 – 281.

[53] M. Thoury, M. Elias, J. M. Frigerio, C. Barthou, Nondestructive varnish identification by ultraviolet fluorescence spectroscopy, Appl. Spectrosc. 61 (2007) 1275–1282.

[54] P. Dondi, L. Lombardi, M. Malagodi, M. Licchelli, T. Rovetta, C. Invernizzi, An interactive tool for speed up the analysis of UV images of Stradivari violins, in: New Trends in Image Analysis and Processing - ICIAP 2015 Workshops, volume 9281 of *Lecture Notes in Computer Science*, Springer International Publishing, 2015, pp. 103–110. doi:10.1007/978-3-319-23222-5_13.

[55] P. Dondi, L. Lombardi, I. Rocca, M. Malagodi, M. Licchelli, Multimodal workflow for the creation of interactive presentations of 360 spin images of historical violins, Multimedia Tools and Applications 77 (2018) 28309–28332.

[56] D. G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comp. Vis. 60 (2004) 91–110.

[57] P. H. Torr, A. Zisserman, Mlesac: A new robust estimator with application to estimating image geometry, Computer vision and image understanding 78 (2000) 138–156.

[58] D. Müllner, Modern hierarchical, agglomerative clustering algorithms, arXiv preprint arXiv:1109.2378 (2011).

[59] D. Arthur, S. Vassilvitskii, k-means++: The advantages of careful seeding, Technical Report, Stanford, 2006.

[60] X. Zhu, C. Change Loy, S. Gong, Constructing robust affinity graphs for spectral clustering, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 1450–1457.

[61] G. J. McLachlan, K. E. Basford, Mixture models: Inference and applications to clustering, volume 38, M. Dekker New York, 1988.

[62] A. Rodriguez, A. Laio, Clustering by fast search and find of density peaks,

science 344 (2014) 1492–1496.
[63] M. J. Rezaee, M. Eshkevari, M. Saberi, O. Hussain, Gbk-means clustering algorithm: An improvement to the k-means algorithm based on the bargaining game, Knowledge-Based Systems 213 (2021) 106672.
[64] Z. Wang, Z. Yu, C. P. Chen, J. You, T. Gu, H.-S. Wong, J. Zhang, Clustering by local gravitation, IEEE transactions on cybernetics 48 (2017) 1383–1396.
[65] R. J. Campello, D. Moulavi, A. Zimek, J. Sander, Hierarchical density estimates for data clustering, visualization, and outlier detection, ACM Transactions on Knowledge Discovery from Data (TKDD) 10 (2015) 1–51.
[66] T. Lei, X. Jia, Y. Zhang, L. He, H. Meng, A. K. Nandi, Significantly fast and robust fuzzy c-means clustering algorithm based on morphological reconstruction and membership filtering, IEEE Transactions on Fuzzy Systems 26 (2018) 3027–3041.
[67] W.-J. Kim, S.-D. Kim, K. Kim, Fast algorithms for binary dilation and erosion using run-length encoding, ETRI journal 27 (2005) 814–817.