Webcam based display calibration

Jean-Baptiste Thomas; University of Burgundy; Dijon, France
Jérémie Gerhardt; Fraunhofer FIRST; Berlin, Germany

Abstract

We present an automatic method for measuring the tone response curve of display devices based on visual methods, where the eye is replaced by an end-user, uncalibrated camera, such as a webcam. Our approach compares a series of halftoned patches of known covering ratio with a continuous series of tone patches for each ratio. Both patches are shot by a camera that is used as a virtual eye to evaluate the luminance difference. By an iterative process, the continuous tone value is adjusted while compared with the perceived level of the halftoned patch. When the camera does not see any difference between the patches or a minimal difference, the luminance level of the continuous patch corresponds to the relative luminance of the halftoned patch covering ratio. We demonstrate that the method is as accurate as an equivalent visual method. The advantage of using a camera over the human eye is due to the limitation of observer variability while performing visual tasks.

Introduction

We are describing an uncalibrated camera based display calibration method. We retrieve the relative luminance response curve of a display device based on visual method, where the eye is replaced by only a camera. The background of our method is to be found in existing methods based on visual calibration approach [1, 2]. The response curve is retrieved by comparison of intensity levels with halftoned color patches. Such methods typically compares a 50% luminance patch and then find a gamma value or compares a series of different halftoned patches with a graylevel patch, before approximating the response curve via an interpolation or approximation processes (typically linear or spline interpolation with or without freedom degrees are used).

Alternative methods augmented with an end-user camera exists [2, 3]. The camera is calibrated relatively to the display by the establishment of a 50% luminance and an offset values, then a pattern of patches from 0 to 1 is displayed and shot by the camera. This method is really sensitive to the black offset estimation and to the 50% luminance evaluation, as demonstrated by Mikalsen et al [3].

Non-visual methods involving camera exist, in the sense of only the intensity signal is measured and no image processing is involved. For example using a spectrometer allows to measure the response curve or by high dynamic range (HDR) imaging. With HDR imaging we can turn a regular a camera in a fairly intensity measuring device[4]. The procedure is known and straighforward to implement [5], but it requires to calibrate the camera. We want to use the most direct approach, therefore we did not re-perform this experiment involving HDR.

Our proposed method starts by the observation that all of the visual methods are supposed to be sensitive to observer differences, either between observers or even between several tries performed by a single observer. We suppose that this variability can be significative, and in order to overcome this problem, we suppose that a camera repeatability would be better than the observer's. We propose thus to use a camera instead of the eye to improve the repeatability of the response curve retrieval.

In order to evaluate the observer variability, a pre-experiment is conducted to evaluate this variation. Then, if the variation intra or inter observers is significantly influencing the result, removing the human subjective response should limit this effect.

Before digging in our camera calibration method, we start this article by a study about observer variability. This experience is run on two groups of observers, which are asked to adjust the intensity level of gray-level patches compared to halftoned patches with a known ratio. Following this purely human observer approach, we will conduct an experimentation using a camera simulating human observer eye for display tone response curve retrieval. Then, results will be compared to other methods such as a totally visual and a spectrometer approaches.

Observer variability

As demonstrated by Mikalsen et al [3], the evaluation of a 50% luminance patch by an observer should not be taken easy, considering that usually, the result of the calibration will higly depends on this evaluation.

The underlying idea behind the proposed method is that observer variation is significant enough, such as shown by Bala et al [2], and that another solution could be considered. We propose in this section to evaluate and quantify the observer variability when asked to perform a luminance matching on displays. We consider two type of variability: inter-observer variations and intra-observer variations.

An experiment is carried out, where observers are asked to adjust the intensity of a patch level compared to a known halftoned level. The task is repeated several times per experiment session (10 times each) for the levels 0.25, 0.5 and 0.75 in a random way. There are three sessions, between each session, the observer is asked to go back to his normal activity or to have a coffee. Figure 1 illustrates the comparison task asked to a panel of observers for the level 0.5.

The experience is performed on two groups of observers, where the luminance levels are adjusted with a slider. Two different displays have been tested, one by each group of people (a group of people consists mostly of end-user is in Germany -Group 1, the other one consists of a mix of experts in imaging fields is in Norway - Group 2): a DLP projector is used for Group 1 in a dark surrounding and an LCD projector is used for the second group in a dim surround.

The results are presented in Table 1 for the global and inter-variability of the two groups. The intra variability is illustrated by observer in Figure 2 and Figure 3 for respectively the first and second group of observers.

Results of the first group show surprisingly a different behaviour than first expected, considering the ability of the visual system to discriminate better the intensity variations in the dark area, following the Weber-Fechner law. We observe that the more repeatable task is actually the one at 50%. The patterns of 25 and 75% being regular. The second group of observers shows interesting results in the way that some observers are consistent in their choice, showing a constant variance. This could come from the habit of being involved in such tasks and in the participation to several psychophysical experiment. This group being composed mostly of experts.

The explaination of the variance is to be found in different aspects:

- First, this can be explained by the adaptation or no adaptation of the observer to this matching task. This could explains the results for the first group, since a 50% luminance matching is a task that a non expert would usually perform for characterization.
- Second, and more likely, this could come from a variation in chromaticity between the halftoned and greylevel patches, especially for the 0.25% patch. It is well known that a chromaticity shift appears in displays while changing luminance level [6]. This task becomes then an heterochromatic luminance matching, which is known to be far more difficult than an homochromatic luminance matching. This fact seems to be confirmed by the answer of the observers to the question *What was the most difficult things for this task?* where most of them answered: *The two patches are not of the same color, which makes it difficult to make them to match*. This is very much the case for the second group as the 3-LCD projector was chowing some misalignment of the red channel for pixels on the right of a black value.
- Last, and not least, there some dithering artefact at the edge between the two patches when other covering ratio than 50% are used. Although, we used the same patterns as described by Neumann et al [1], but these patterns show an artefact anyway. This is a major issue in the success of such a task as the observer aims usually at making this edge to disappear.

Table 1 tells us that the variation in digital value, considering the standard deviation of the successive attempts can be up to 10 digits (on a typical 8 bit device). We then confirm the variability observed by Bala's work [2], where they did observe some variability on the computed gamma values based on a 50% patch (observation confirmed by Mikalsen [3] while evaluating Bala's method).

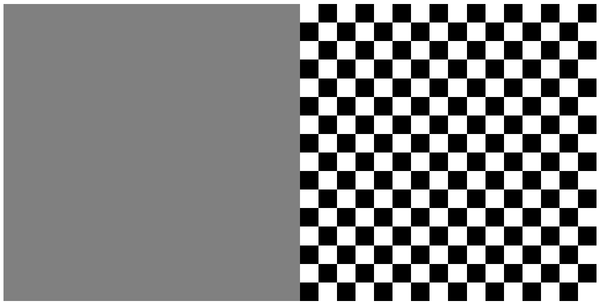

Figure 1. *Illustration of the patch comparison performed both visually and automatically with the camera. The right halftoned patch has a known ratio white versus black pixel when the left patch has a single continuous tone value, which needs to be adjusted.*

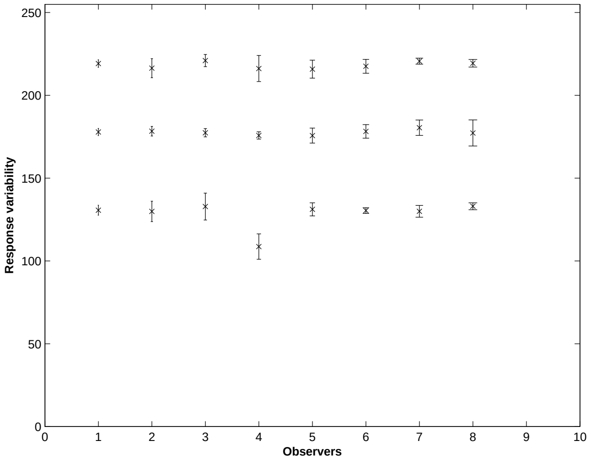

Figure 2. *The graph illustrates the intra variability per observer and per level of the first group. We plot the standard deviation centered on the mean value.*

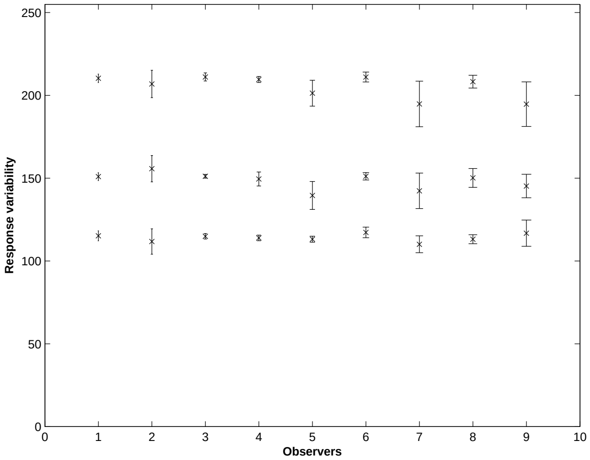

Figure 3. *The graph illustrates the intra variability per observer and per level of the second group. We plot the standard deviation centered on the mean value.*

Observer free evaluation

In this section, we replace the observer's eye by an uncalibrated camera. The idea is to perform a similar observer task than in Neumann et al. [1] paper, but the matching should be assessed automatically by the camera. For this purpose, a set of halftoned matching patterns are displayed and compared with an adjusted intensity value.

The interest of developing such a process is double: first, being free from observer variability, secondly potentially computing a large number of data without any boring effect due to a repetitive visual task. Other advantages include that, if a webcam is

	Data 1			Data 2		
level	0.25	0.5	0.75	0.25	0.5	0.75
ci μ_{high}	0.5007	0.6955	0.8546	0.4484	0.5840	0.8077
μ	0.5031	0.6966	0.8559	0.4471	0.5821	0.8053
ci μ_{low}	0.5055	0.6978	0.8572	0.4459	0.5801	0.8028
ci σ_{high}	0.0340	0.0168	0.0185	0.0203	0.0323	0.0407
σ	0.0356	0.0176	0.0193	0.0194	0.0308	0.0389
ci σ_{low}	0.0374	0.0184	0.0203	0.0186	0.0295	0.0372
$\sigma * 255$	9.07	4.48	4.93	4.95	7.85	9.92

Inter variability between the two groups of observers (Data 1 and 2). Average (μ), standard deviation ($sigma$) and confidence interval (ci%) high and low for both average and standard deviation are displayed for the three levels 0.25, 0.5 and 0.75 that were asked to be adjusted for the intensity level comparing with an halftoned patch of the corresponding recovering ratio. The first group shows a bad evaluation in the dark area, where there could be some chromaticity shift, while the second group follows the rule that higher luminance are more difficult to discreminate.

used, one do not need to take care of the camera settings: each evaluation is independent and the camera settings are automatically setup in order to have the best discrimination (considering that the camera does not do any image subjective content based processing).

Camera consideration

The camera should be registered geometrically and images corrected by a projective transform. The camera should also acquire the halftoned pattern as uniform flat color patches. We thus could include a blurring effect if needed. The blurring effect is realized by optically defocusing either the camera or the projector after camera registration. An optical blurr is a linear transform and does not imply problems with the uncalibrated camera, at the difference of a digital blurring filter apply on the acquired image, which would imply inaccuracies due to camera potential non linearities. An optical defocus could induce some chromatic aberrations, but we did not observed this.

While comparing the halftoned patch and the continuous patch for checking the luminance, we assume the display to have chromaticity consistency. This assumption is shown to be critically wrong at some point and can be a limitation of the method as the spectral properties of the camera are fixed.

Last, there could be some issues with the camera considering the variation between pixels and the spatial non uniformity due to the lens or other abberration effects. This can be taken into account with to simple things: taking into account only a small area at the center of the camera, and displaying the halftoned patch and the continuous patches at the same location, then averaging the pixel values. Note that, while doing this, we take into account the spatial non-uniformity of the display also.

We make the hypothesis that the chromaticity shift of the projector while going toward dark luminances [6] will not affect too much the camera luminance matching. This is not obvious as camera will not change its transmittance filters and they are different from the human eye (and different of a standard luminance spectral curve).

Luminance adjustment - algorithm

The color patch (left patch in Figure 1) should be then modified (manually or automaticaly) until the variation with the halftoned patch (right patch in Figure 1) disappears or becomes as small as possible from the camera point of view. In order to do that, we compute the Euclidean distance between the camera digital values.

We considered several approaches to adjust the intensity level versus a known halftoned level:

- Comparing the two patches difference with a Δ tolerance limit and increasing or decreasing the adjusted value until the condition is satisfied. This solution is somehow not so robust since the tolerance value need to be carefully defined and a good value for low intensities can be not very good for high intensities. The tolerance value could also follow a function of the intensity, in order to be modulated. This solution did not give very robust results, and we abandonned it.
- Following the same principle, but setting up an optimization process did not give good results, because the camera changes settings between different patches, thus create some local minima, which leads to a lack of robustness.
- Computing the differences for a set of input values and approximate the function $\Delta = f(luminanceLevelCamera, displayDigitalValueCamera)$, where $luminanceLevelCamera$ is the halftoned patch of a given covering percent -luminance value- shot by the camera and $displayDigitalValueCamera$ is the continuous color patch from a given digital value shot by the camera. Δ is the euclidean distance between the camera digital values. Then, find the minimum of this function, and so find the best digital value that correspond to the intensity value. In doing this, we avoid the problem with local minima (we find the annulation of the local derivative, but being sure that it is not a local minima), we get rid of the tolerance value. An example of the function is shown in Figure 4. The discretization of this function is shown to be of major importance in the following. As shown in Figure 5 and Table 2, if the discretization step of this function (in other words, the number of shot required to evaluate f) is too small, the function itself could be the same for different halftoned patches and then leads to the same minimum for different intensity values.

The pseudo-algortihm used is written below.

Discretization	0.02	0.05	0.1	0.2	0.3	0.4	0.5	*Visual*
Error	0.0581	0.0586	0.0588	0.0748	0.0848	0.1328	0.1533	0.0521

Relative error from the estimation of the curve of the 3-LCD projector and the spectrometer reference at 10 points for a different discretization step and for the visual assesment. When the discretization step decreases, the results tend to be as good as an observer evaluation.

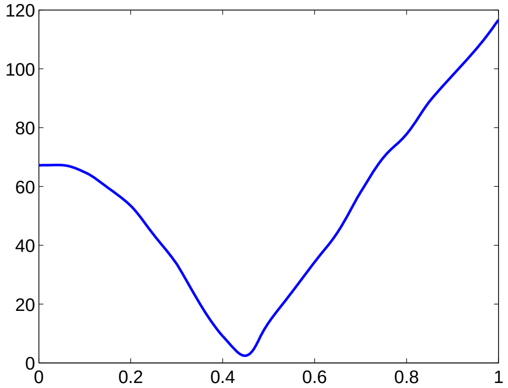

Figure 4. *A typical shape for function f. This function minimum is used to assess the luminance matching between an halftoned pattern and a projector gray level.*

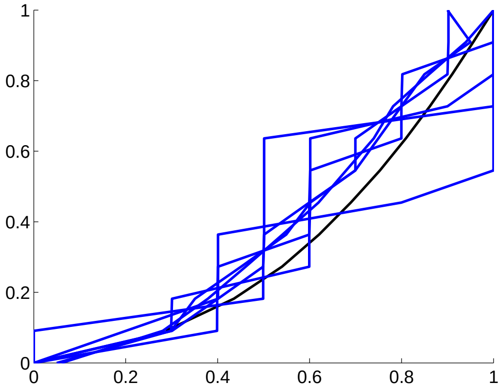

Figure 5. *Response curve of a 3-LCD video-projector based on 10 halftoned patches with different discretization steps for evaluate f: 0.02, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5 has been taken on a normalized scale. We can see that the discretization influence the selection of the right level: This is due to our method -evaluation of f, then find its minimum-. The reduction of the digitization step is done at the expense of the number of shot taken by the camera, and then of the time to evaluate the response curve. A correct value can vary from a device to another.*

Results and discussion

The experiment was performed on two projection systems with a webcam. The first setup was in a dark room with a DLP projector. The second setup was in a dim room with a 3-LCD projector. The cameras were two different Logitech Webcam Pro 9000. We decided to vary the condition and material to demonstrate the robustness.

Algorithm 1 TRC estimation; N is the number of luminance level evaluated, M is the number of continuous patches used to evaluate the function f - f is as defined above in the text.

```
procedure ESTIMATETRC(in N, M, out $dataTRC[]$)
    $p \leftarrow \frac{1}{N}$
    $q \leftarrow \frac{1}{M}$
    $Initialize\ dataTRC[]$
    for $i \leftarrow 0$ to $i < 1$ with $i \leftarrow i + p$ do
        $digitalValueEq \leftarrow 0$
        $Initialize\ tabDef[]$
        $luminanceLevelCamera \leftarrow shootPattern()$
        for $j \leftarrow 0$ to $j < 1$ with $j \leftarrow i + q$ do
            $displayDigitalValueCamera \leftarrow shootPattern()$
            $tabDef[j] \leftarrow$ $f(luminanceLevelCamera, displayDigitalValueCamera)$
        end for
        $f \leftarrow estimateF(tabDef)$
        $digitalValueEq \leftarrow f^{-1}(min(f))$
        $dataTRC[i] \leftarrow digitalValueEq$
    end for
end procedure
```

The Figures 6 and 7 show the response curve evaluation by three different methods for each projector: by visual assessment, automatic camera measurement and spectrometer. The visual assessment is likely the same than the test performed to evaluate the observer variability but with more halftoned covering percentage steps. The automatic camera measurement mimics the eye processing by comparing an halftoned patch and a continuous tone patch as explained above. Both methods evaluate 10 luminance steps linearly distributed in the digital display steps. Finally the spectrometer measures the light projected by the projector at various ramp steps.

We remind that we want to avoid the observer feedback, which introduces a certain variability in the results. Our automatic approach corrects this observer dependence of the method and therefore increase its repeatability. We ran several times our method and observed similar results, thus we are confident that our method is robust enough for practical use. This is illustrated in Figure 8, where the evaluation has been ran two times with the same parameters, giving similar results.

We can notice that we are having better or similar results than the observer in the dark areas. We thought originally that increasing the number of data in this area in order to fit even better the actual response curve would be of benefit, however as shown in Figures 8 and 9 where we increased the number of patches to 100 and reduce the discretization step to 0.01, the result is still different from the reference measurement. We have thought first that the black level was of great influence and considered to correct the data with a linear transformation such as $LuminanceCorrected = (1 - \%halftoned) * blackLevel + \%halftoned$, but this would simply push the low luminance data to the level of the reference curve, but would not change a lot the values of the upper luminances, thus a black level correction is not enough to explain this fact. An explanation could be found while looking toward color/luminance appearance models, as the user and camera choice could be influenced by the background of the target, through some glare effect for the camera and some simultaneous contrast and assimilation effect for the observer. This require further work to be demonstrated.

We observe also some variability in the high luminances on both set-ups. It is surely due to the camera automatic setup and the reach of a saturation intensity. The camera will show some limitations toward a value where it would saturate.

The precision of our approach is subject to the number of steps while the algorithm is adjusting the intensity level as shown in Table 2 anf the corresponding Figure 5. The jumps in the displayed response curve (see Figure 6) for both the visual and automatic methods in comparison to the spectrometer approach highlight the importance of the fitting method used after measurements.

To conclude with these results, our automatic approach gives good and stable results. The fact that there is a chromaticity shift of the pattern does not seem to be a problem for the camera method to give equivalent results than the observer, which was one of the hypothesis we made.

Conclusion

This work shows that we could use a webcam to replace the human eye in some visual assessment for display calibration. We demonstrated that it is possible to retrieve the luminance response curve of a display with such a method. Our method gives equivalent results than while performing the visual task directly, with all the convenience of automatic methods.

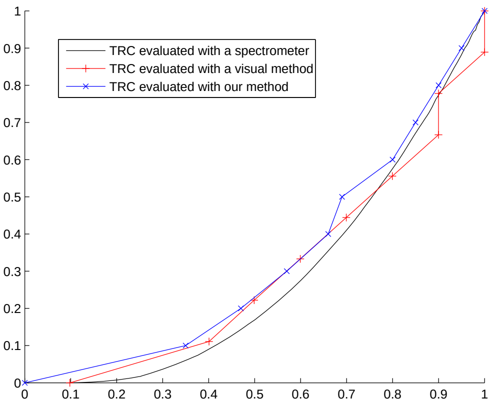

Figure 6. *Tone response curve evaluation for the DLP projector by three different methods: Visual, webcam and spectrometer (the reference). We can see that our method is working pretty well with this set-up. The chosen discretization step was 0.1 for the evaluation of f. The results are similar to the observers.*

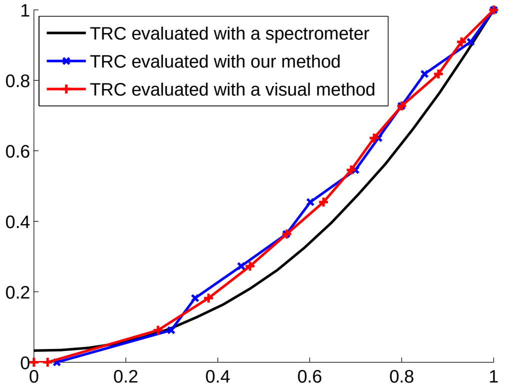

Figure 7. *Tone response curve evaluation for the 3-LCD projector by three different methods: Visual, webcam and spectrometer (the reference). We can see that our method is working pretty well with this set-up. The chosen discretization step was 0.05 for the evaluation of f. The results are similar to the observers.*

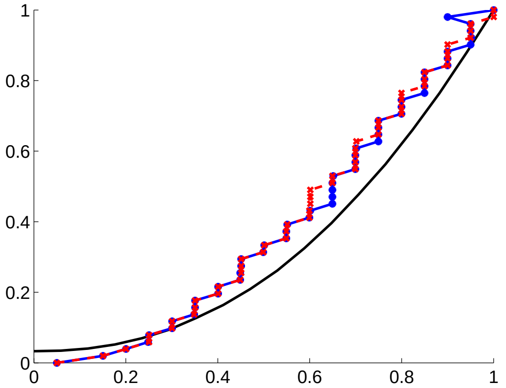

Figure 8. *Two evaluation of the response curve of the 3-LCD projector with 50 luminance values and a discretization step of 0.05; the results are relatively consistents.*

Acknowledgments

We would like to thank the persons who participated in the subjective evaluation of luminance patches, we know it is quite a boring task. We also thank our respective institutions for financial support.

References

[1] Attila Neumann and Alessandro Artusi and Georg Zotti and Laszlo Neumann and Werner Purgathofer, An Interactive Perception Based Model for Characterization of Display Devices, Proc. Color Imaging Conference, pg. 666. (2004).

[2] Raja Bala and Karen Braun, Efficient and simple methods for display tone-response characterization, J. of the Society for Information Display, 15, 11, (2007).

[3] Espen Bårdsnes Mikalsen and Jon Yngve Hardeberg and Jean-Baptiste Thomas, Verification and extension of a camera-based end-user calibration method for projection displays, Proc. of CGIV 2008 - Fourth European Conference on Color in Graphics, Imaging and

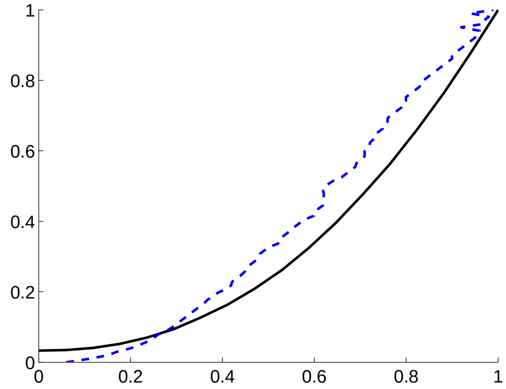

Figure 9. *Evaluation of the response curve of the 3-LCD projector with 100 patches with a discretization step of 0.01. The method is relatively stable from values close to each other. We can notice some unconstancy on the high luminances.*

Vision, (2008).

[4] Jérémie Gerhardt, Simple HDR approach for projector spatial nonuniformity correction, Proc. Worshop Farbbildverarbeitung, Berlin (2009).

[5] Paul E. Debevek and Jitendra Malik, Recovering high dynamic range radiance maps from photographs, Proc. of SIGGRAPH 97: Proc of the 24th annual conference on Computer graphics and interactive techniques, pg. 369-378 (1997).

[6] Jean-Baptiste Thomas, Jon Y. Hardeberg, Irene Foucherot and Pierre Gouton, The PLVC color characterization model revisited, Color Research & Application, 33, 6, pp 449-460, (2008).

Author Biography

Jean-Baptiste Thomas received his Bachelor in Applied Physics in 2004 and his Master in Optics, Image and Vision in 2006, both from the Université Jean Monnet in France. He received his PhD from the Université of Bourgogne in 2009. After a stay as a researcher at the Gjøvik University College and then at the C2RMF (Centre de Recherche et de Restauration des Musées de France), he is now Maître de Conférences (Associate Professor) at the Université of Bourgogne. His research focuses on color science and on color and multi-spectral imaging.

Jérémie Gerhardt received his Master in image processing from the University Pierre et Marie Curie (2002) and his PhD in image and signal processing from the Ecole Nationale Supérieure des Télécommunications (ENST) in Paris (2007) but did his research at the Norwegian Color Research Laboratory in Gjovik in Norway. Since then he has worked in the Fraunhofer FIRST (2008) in Berlin, Germany. His research work took place in the field of spectral color imaging, in his beginning on spectral color reproduction with multi-ink printer and yet on color correction for tiled displays installation and image processing.