On the Selection of Patches for Color Camera Calibration

Margarita Khokhlova, Jon Yngve Hardeberg; The Norwegian Color and Visual Computing Laboratory, Gjøvik University College; Gjøvik, Norway.

Abstract

This research studies the influence of the color model used for the optimal sample selection from a larger dataset, for the task of color camera calibration. Most algorithms proposed in the literature perform such sample selection in the CIELAB color space, however since color transformations from one space to another are non-linear, choice of color space will affect the final target design. This work establishes the relationship between the color space where the patches are selected and the accuracy of the final calibration result. The Kennard-Stone algorithm with different distance metrics is used to choose a fixed number of patches for the calibration target. Final calibration results are compared for various selections of patches from three datasets, using different state of the art color camera calibration methods. The results are also compared to those obtained with a classical Macbeth Color Checker. This research highlights the importance of setting proper parameters for color patches selection for custom calibration target design.

Introduction

Accurate color camera calibration is a critical task for many applications where reliable color matching is needed. Color camera calibration here refers to defining and applying a correction function for the transformation from the device-dependent camera response color space to a device-independent standard color space. To calibrate a camera one typically needs the standard color space values of a reference image, and experimental observations of how the camera interprets this reference. Usually, a calibration target with known color characteristics is used to estimate this relationship between the RGB camera output values and the known reference color values.

Existing color targets vary depending on the application but are typically general-purpose in nature. Therefore, using such standard color targets does not guarantee a good basis of training samples for camera calibration, because these samples might not be optimal for a given specific application general perspective [6].

In addition, many existing standard color targets contain quite large amounts of colors (x-rite ColorChecker Digital SG: 140, IT8.7/2: 288, Datacolor SpyderCHECKR: 48 etc.). Such standard color targets could be inconvenient or even impossible to use in some applications. Sometimes a custom color target is required, allowing for designing an optimal training set for the application, with appropriate number and properties of patches, and more accurate calibration can result. Such custom color calibration target patch selection is the subject of this research. After describing related state of the art in the next section, we present the proposed methodology of this paper, followed by the experimental setup and results, before concluding and identifying areas of further research.

Background

Several methods [1-5] for creating high precision color targets by selecting samples from a larger datasets have been described in the literature. Andersen and Hardeberg [1] proposed a method of choosing the most significant patches from a set based on camera hue domain subdivision. Samples with the lowest susceptibility to noise are selected in all hue directions in a way so they do not overlap each other in colorimetric hue correlate. The final subset design helps to make camera calibration preserving an original white point value. The Kennard-Stone (KS) [5] design aims at selecting samples normally distributed within the given set. The Kang (KG) method [3] first divides the unlabeled color patches into clusters by similarity and then selects from each cluster the most representative example. The most representative example in their work is the one closest to the cluster's centroid. These selected representative patches are then used for color calibration. Pellegri et al. [2] compare many existing patch selection algorithms and propose three new approaches for selecting a training set to be used for the color characterization of a multispectral acquisition system. The methods are Hue Analysis method based on colorimetric considerations, Camera Output Analysis method and Linear Distance Maximization method based on algebraic and geometrical facts. They also evaluate the influence of the number of samples in a target design on calibration accuracy. Eckhard et al. [6] compare the performance of many existing algorithms and conclude that some methods are better than the others depending on the size of the initial dataset. They also propose their own method called Recursive Rejection (RR) which iteratively removes samples from the dataset. The RR method also starts with a clustering as [3] but then it optimizes a training set for a specific application by measuring the change in performance in spectral estimation when certain clusters of samples are rejected from the training. Alsam and Finlayson [7] introduce an optimal sample reduction algorithm based on integer programming, which is driven by camera colorimetric and spectral calibration. Zhang et al. [4] propose two methods for the selection of representative color samples for the spectral digital device characterization. In contrast to [2] and [5], researchers try to select not the most distinct samples from a dataset, but the most representative ones. The first method selects the training samples whose eigenvectors can accurately model the reflectance set. The second method attempts to find in the set of training samples that minimize the difference between the actual and estimated reflectances of a virtual-imaging system.

In some works mentioned, researchers select samples in a camera color space [1] or from spectral data [7], but mostly researchers [2-4,6,19] work in a device-independent color space, preferring to convert to CIELAB (or similarly CIELUV) color spaces before running their patch selection algorithms. The main reason to do so is the fact that the distances between colors in CIELAB color space match the color differences for a human observer, measured with the ΔE^*_{ab} standardized color difference. However, the transformation from a camera color space to a device-independent space is not linear. This means that the final design of a color target might vary depending on the color model used for the selection procedure. Then there is reason to suspect that the resulting calibration designs obtained using the CIELAB color model are not necessarily optimal.

Proposed methodology

Camera Simulation

For the task of a spectral characterization of an electronic camera, it is necessary to choose a model of the system. In this work camera simulation is performed. For the sake of simplicity, an ideal model without noise is used. The assumed camera sensitivity function showed on Figure 1.

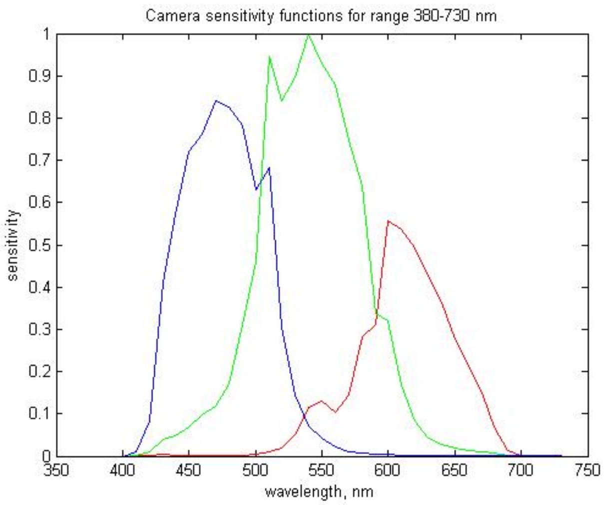

Figure 1. *Sensitivity function of the camera model used in simulations.*

The model of camera selected corresponds to sensitivity functions of the industrial camera which is a subject of an interest for this work. No noise were introduced to the camera model since the main goal is to check the influence of the color space, where samples are selected, so as long the model is consistent for each experiment, it is enough for the goal of this research.

Color camera calibration

As stated before, color calibration could be viewed as modeling the relationship between the input and the output of the camera. Selected color patches characteristics are used to build either a set of linear or polynomial equations used for regression analysis. Color camera calibration could be performed using techniques such as linear or polynomial regression. The easiest way to perform a camera calibration is the one uses Moore-Penrose pseudo-inverse matrix [9]. The color correction linear transform is then just a 3×3 matrix A which performs a mapping of a patch camera color value RGB into the $CIEXYZ$ space:

$$XYZ = RGB \times A, \tag{1}$$

in which case, the camera calibration matrix estimation reduces to the calculation:

$$A = RGB^{+}_{(3x10)} \times XYZ_{(3x10)}, \tag{2}$$

where RGB^{+} denotes the Moore-Penrose pseudo-inverse. In the case oflinear regression, even a small number of patches is enough to perform the camera calibration, because A has eight degrees of freedom. We choose the minimum number of colors in our target to be 10.

Polynomial regression is another approach to estimate the color correction transform which usually gives a better fitting with a lower mean error value [10]. A transformation function then can be equivalent to the following equation for the second-order polynomial [11]:

$$A = [1\ R\ G\ B\ R^2\ G^2\ B^2\ RG\ GB\ RB\ RGB], \tag{3}$$

or to the third-order polynomial regression as in [9]:

$$A = [1\ R\ G\ B\ R^2\ RG\ G^2\ GB\ B^2\ R^3\ R^2G\ R^2B\ RG^2\ RB\ RB^2\ G^3\ G^2B\ GB^2\ B^3] \tag{4}$$

A very important factor concerning the success of the polynomial regression algorithm is the choice of the functions defining the vectors [8]. Testing other functions and, especially, third-order polynomial functions, which give a better fit, is important to obtain the best results. However, on practice linear regression is often chosen for camera calibration for the case of simplicity [12]. In this work we experimented with both linear and polynomial regression for color camera calibration.

Color space transformations

For this work, color patches are selected from the datasets based on their coordinates in different color spaces. In total four standard color spaces were used: sRGB, CIEXYZ, CIELAB and CIELUV. Color information in the datasets is presented in the form of spectral data in range 380-730 nm, at 10 nm intervals. All the transformations between color spaces are done using standard color science formulas [13].

Datasets

Three different datasets were used to perform the experiment. The datasets are composed of different materials to generalize the results of this research. The first dataset is the 1269 samples of the Munsell Book of Color available from the University of Eastern Finland [14]. The second dataset is a set of Textile color samples consisting of 4826 samples of different fabrics, courtesy of Professor Seyed Hossein Amirshahi, Amirkabir University of Technology, Iran. The third dataset consists of 3818 measured reflectances of printed color samples measured at Gjøvik University College [15].

Set of parameters

The main interest of our experiment is to establish the influence of a color space used on calibration results for an optimal sample set. However, we extend this goal by also including the evaluation of the performance of different distance metrics and calibration methods. Commonly [2][3], the Euclidean distance is chosen as a distance metric because in CIELAB space it corresponds quite well to perceptual color differences. In this work also the CityBlock and CIEDE2000 [16] distances are tested. Using different types of distances is not a new approach, for example, Pellegri et al. [2] test their algorithms using the Euclidean distance and the maximum absolute distance of coordinates as distance metrics. However, we propose a more complete data set in our experiment. We vary the distance, the color space and the calibration method for the 3 different datasets.

Experimental setup

Several experiments were conducted to establish the optimal color space for samples selection. The target design which consists of a given number of distinct points is chosen among the candidates points of 3 datasets with the method proposed by Kennard and Stone [5]. This algorithm is very simple and general, so we assume that the results obtained could be also implied for the other methods based on a color distance between points such as [2][3]. The final subset of points is chosen sequentially. First two points in the design are two farthest away points in the dataset point cloud.

At each following iteration of the algorithm, the aim is to have the points in the subset "uniformly" spaced over the color space, according to the distance metric used:

$$D = dist(x_v, x_m), \quad (5)$$

where x_v is the set of selected points, x_m are the candidate points, and dist the distance metric used.

Then the maximum distance between candidate points and the nearest points in the dataset is calculated and the new point is added to the subset. Thus, the algorithm chooses the point among those remaining that is farthest from an existing design point. Example for the patch selection with the Kennard-Stone algorithm is shown in Figure 2.

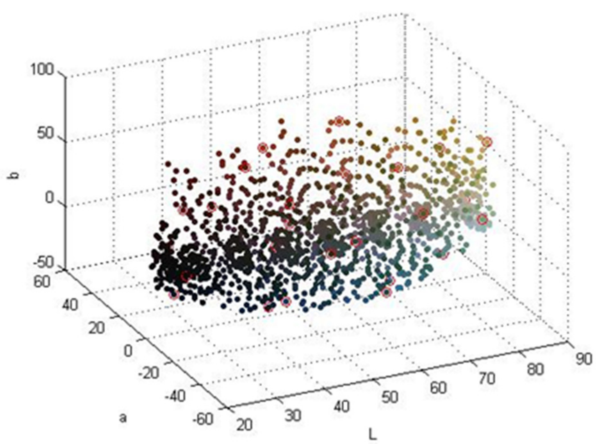

Figure 2. Optimal patches selected (in red) for the Munsell dataset [6] using the Kennard-Stone method in the CIELAB color space.

A detailed description of the method could be found in the paper published by Kennard and Stone [5]. The algorithm tends to show good results for general case applications [2][6]. The Kennard-Stone algorithm does not have a single convergence factor. The stopping criteria used in this work is the number of samples in the final set, so we select sequentially N samples.

The correction function is then obtained using the designed color target and formulas 1-4. We also tested regression not only from sRGB to CIEXYZ color space, but also regression from sRGB to CIELAB color space. In the last case cubic root function is applied on sRGB values before the regression as in [8].

The errors of performed color camera calibration are evaluated using the standard distance metric ΔE^*_{ab}. We evaluate the error in terms of mean and max ΔE^*_{ab} values for a given dataset and the selected color target itself. This metric is widely adopted for color camera calibration evaluation [1][2][7] and it gives an opportunity to evaluate results in terms of Just Noticeable Difference (JND) [17]. In order to show that a custom color selection for a color camera calibration is preferred to a standard one, we also obtain the calibration functions using the Macbeth ColorChecker color patches [18].

Results

We conduct series of experiments in order to estimate the influence of different parameters on the final target design and calibration results. Using the Kennard-Stone method for patches selection, we modify patches selection and camera calibration parameters in order to establish the relationship between them.

Distance metrics

In the first trial of experiments, we use different distance metrics to select patches: the Euclidean distance, the CityBlock distance and the CIEDE2000 distance (the last was used only for selection in the CIELAB color model). The camera is then calibrated with the obtained color targets and the ΔE^*_{ab} error is calculated. The experiment shows that the Euclidean distance metric between coordinates gives best results and only in one case using CIEDE 2000 gives slightly smaller mean ΔE^*_{ab} error. The CityBlock distance either gives exactly the same results as the Euclidean distance if the tested dataset is not very big and does not contain many patches with similar coordinates, either just slightly different results. With the results obtained for all 3 datasets, the Euclidean distance proved to be the best one for patches selection using the Kennard-Stone Method, no matter in which color space the selection is performed.

Comparison to the custom color target

The next step is to compare the custom color targets designed of 24 colors to the Macbeth ColorChecker standard one. As mentioned before, the Macbeth ColorChecker is widely used for a goal of color calibration of digital devices. This test is critical for our experiment since we claim that it is better to make a custom color target for a dataset than to use a standard one. The outcome of the test proves that custom target creation method gives better results in terms of mean ΔE^*_{ab} error for two out of three the datasets tested. Only for the Munsell dataset, the calibration using the Macbeth ColorChecker chart gives slightly better result in terms of ΔE^*_{ab} mean error. In this trial, we designed our targets of 24 colors to allow a comparison with the standard color target which contains 24 colors. The result obtained corresponds quite well for the fact that Munsell Dataset is a classical one, so it was probably covered when creating a Macbeth ColorChecker. The other two datasets are more specific and contain more different colors. But it is important to mention that while the difference between mean error obtained with custom and standard targets for Munsell dataset is not big, for the other two the custom patches selection gives a very big improvement in terms of mean and max ΔE*ab error values. The results are summarized in Tables 1-2 for the Textile set.

Camera calibration method

Evaluation of the calibration method selection on final calibration results is the goal of the next trial of experiments. In these we perform the color camera calibration using four different regression techniques mentioned previously: linear regression from sRGB to CIEXYZ space, linear regression from sRGB to CIELAB space, polynomial regression from sRGB to CIEXYZ and polynomial regression from sRGB to CIELAB color space. Results obtained show that polynomial regression from sRGB to CIELAB color space outperforms other calibration techniques. The third-order polynomial function 4 gives the best results. Table 3 shows the results obtained by the calibration with the 3-rd order polynomial regression for the HIG dataset. Overall, even with the 3-rd order polynomial regression we have quite big mean and max ΔE*ab error values, but this is due to the small number of patches in the final target design.

This corresponds well to the observations made in other work in this field [8]. The exception in our experiment is only the HIG dataset where better results in terms of mean ΔE*ab error are obtained with the polynomial regression from sRGB to CIEXYZ color space. However, in this case, the differences between mean ΔE*ab error are not very big, but the max ΔE*ab error is way bigger for sRGB-CIEXYZ polynomial regression than sRGB- CIELAB polynomial regression.

Table 1. Calibration results for the Textile Dataset using the Macbeth ColorChecker color target

Regression model	ΔE mean target	ΔE mean set	ΔE max set
linear RGB-XYZ	3.69	4.84	30.96
linear RGB-LAB	6.66	7.11	27.90
polynomial RGB-XYZ	2.85	7.63	109.69
polynomial RGB-LAB	2.07	4.47	15.41

Optimal space for the calibration patches selection

CIELAB space is proved to give good results in general and the optimal color target design is obtained selecting patches from sample points in CIELAB space. It is still not the case for all the calibration methods tested. For example, for the linear mapping which is widely used for camera calibration, the optimal minimum set in terms of mean ΔE*ab error was obtained in sRGB space for all datasets tested. A linear mapping is widely used for its' simplicity and due to the fact that it requires fewer color samples in the final design than polynomial regression methods.

Small color target composition

In the next series of experiments, we design color targets composed of just 10 colors. 10 colors are sufficient enough only for the linear calibration method since the matrix A in formula 1 has 8 degrees of freedom. It is impossible to use just 10 patches for the polynomial regression based calibration methods. However, there are cases when the simplicity and speed are more important than the more accurate results. Once again, the smallest mean error is obtained while selecting patches in sRGB color space rather than CIELAB one using the Euclidean distance metric. An interesting observation is that the better results for all three datasets are obtained with linear regression from sRGB to CIEXYZ color space and not with the regression from sRGB to CIELAB color space. The max ΔE*ab error value, in contrast, tends to be much smaller with the regression from sRGB to CIELAB. Overall the mean and max ΔE*ab errors proved to be bigger using linear regression than using the polynomial regression methods.

Table 2. Calibration results for the custom designed 24 color target for the Textile Dataset.

color space	Dist	ΔE mean target	ΔE mean set	ΔE max set
Polynomial RGB-XYZ, K=24				
RGB	Euclidean	1.86	2.77	13.75
XYZ	Euclidean	1.60	2.56	37.50
LAB	Euclidean	2.53	2.42	18.12
LAB	CIEDE2000	2.83	2.47	12.22
LUV	Euclidean	2.20	2.86	23.31
Linear RGB-XYZ, K=24				
RGB	Euclidean	3.64	2.79	26.49
XYZ	Euclidean	2.08	2.89	48.79
LAB	Euclidean	3.61	2.99	35.91
LAB	CIEDE2000	4.43	2.87	32.88
LUV	Euclidean	3.66	2.74	35.91
RGB-LAB, linear, K=24				
RGB	Euclidean	5.24	5.11	14.18
XYZ	Euclidean	4.56	5.15	19.63
LAB	Euclidean	6.10	5.01	15.98
LAB	CIEDE2000	6.19	4.88	14.43
LUV	Euclidean	5.37	4.81	16.83
RGB-LAB, polynomial, K=24				
RGB	Euclidean	1.32	2.44	14.06
XYZ	Euclidean	1.34	2.29	17.85
LAB	Euclidean	2.13	1.95	11.31
LAB	CIEDE2000	1.96	1.90	9.99
LUV	Euclidean	1.60	2.62	23.79

Table 3. Calibration results for the custom designed 25 color target for the HIG Dataset.

color space	ΔE target	ΔE mean set	ΔE max set
RGB	2.87	2.87	14.84
XYZ	0.95	2.76	19.63
CIELAB	1.13	1.97	19.55
CIELUV	3.72	4.18	20.63

Optimal number of patches in the target

To finish the experiments and establish the optimal number of patches in the target design, we also compute the mean error value function depending on the number of patches used in a color target. It is in not a novel experiment, since the influence of the number of patches was explored before in [6], but it helps to construct an optimal set in case when no constraints on number of patches apply. For two out of three datasets, 50 samples are enough for the optimal color target design and further increase of the number of color patches doesn't give any significant improvement. This result is slightly different from the one obtained in [20], where authors claim that 80 samples are enough to account fully for the Munsell Book of Color, however we use other selection methods. For the test, polynomial regression technique was chosen because it proved to be most accurate in the previous experiments. The result obtained with the Munsell Dataset is shown in Figure 3.

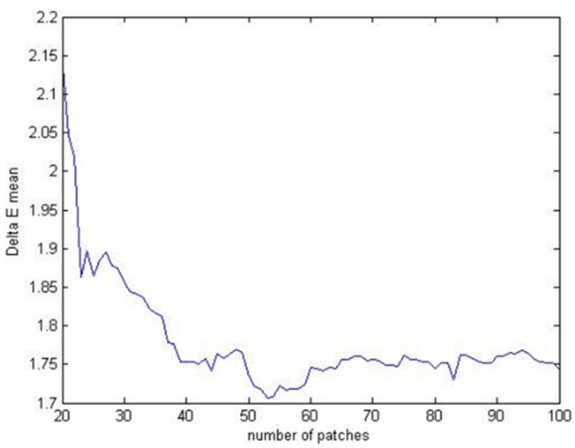

Figure 3. *The mean ΔE*ab error for increasing number of the color patches in the calibration target design. Munsell Dataset.*

Conclusion and Future work

The research presented in this paper shows that the choice of color space in which to perform a sample selection algorithm, in conjunction with the calibration method, is of a great importance for the final color calibration results. To obtain the most accurate results for a given application, a camera should be calibrated with a custom color target, and an appropriate calibration method should be used. Our experiment shows that the creation of a target design is not a straightforward task and depends not only on the sample selection method but also on the color space and calibration method used.

For future work, in addition to extending the experimental basis with more datasets, it would be interesting to explore the use of other color difference equations. This is expected to be especially useful for cases when the colors in the dataset are not normally distributed.

References

[1] C.F. Andersen, J.Y. Hardeberg. Colorimetric Characterization of Digital Cameras Preserving Hue Planes. 13th Color and Imaging Conference Final Program and Proceedings, Society for Imaging Science and Technology, 141-146, 2005.

[2] P. Pellegri, G. Novati, R. Schettini, Training set selection for multispectral imaging systems characterization, *Journal of Imaging Science and Technology*, 48(3):203-210, 2004

[3] H. R. Kang, Color scanner calibration, *Journal of Imaging Science and Technology*, 36(2):162-170, 1992.

[4] H. Shen, H. Zhang, J. Xin, S. Shao, Optimal selection of representative colors for spectral reflectance reconstruction in a multispectral imaging system, *Applied Optics*, 47(13):2494-2502, 2008

[5] R. Kennard, L. Stone. Computer aided design of experiments, *Technometrics*, 11(1):137-148, 1969

[6] T. Eckhard, E.M. Valero, J.Hernández-Andrés, M. Schnitzlein, Adaptive global training set selection for spectral estimation of printed inks using reflectance modeling, *Applied Optics*, 53(4):709-719, 2014

[7] Alsam, Ali, and Graham Finlayson. Integer programming for optimal reduction of calibration targets, *Color Research & Application,* 33(3): 212-220, 2008

[8] J.Y. Hardeberg, Colorimetric scanner characterization, *Acta Graphica,* 155, 2005

[9] A. Albert, Regression and the Moore-Penrose pseudoinverse, Academic Press, New York, NY, 1972

[10] J.Y. Hardeberg, Acquisition and reproduction of color images: colorimetric and multispectral approaches, Universal Publishers, Parkland, FL, 2001

[11] A. Khandual, G. Baciu, J. Hu, D. Zheng, Color Characterization for Scanners: Dpi and Color Coordinate Issues, *International Journal of Advanced Research in Computer Science and Software Engineering*, 2(19), 2012

[12] R.S. Berns, Billmeyer and Saltzman's Principles of Color Technology, Wiley, New York, NY, 2000

[13] Colorimetry - Part 3: CIE Tristimulus Values, Joint ISO/CIE Standard ISO 11664-3:2012(E)/CIE S 014-3/E:2011.

[14] J. Orava. The reflectance spectra of 1600 glossy Munsell color chips. University of Eastern Finland. Online: http://www.uef.fi/fi/spectral/munsell-colors-glossy-all-spectrofotometer-measured

[15] J.T. Simon-Liedtke, J.Y. Hardeberg., P.O. Husøy, Pixel-wise illuminant estimation for mixed illuminant scenes based on near-infrared camera information, Color and Imaging Conference, IS&T, 217-221, 2014

[16] G. Sharma, W. Wu, E.N. Dalal, The CIEDE2000 Color-Difference Formula: Implementation Notes, Supplementary Test Data, and Mathematical Observations, Color Research & Application, 30(1):21-30, 2004

[17] D.L. MacAdam, Visual sensitivities to color differences in daylight, Journal of the Optical Society of America, 32(5):247-273, 1942

[18] The ColorChecker. ColorChecker data (RGB coordinates, spectral data, Excel tables). Online: http://www.babelcolor.com/main_level/ColorChecker.htm

[19] H. Wannous, Y. Lucas, S. Treuillet, A. Mansouri, Y. Voisin, Improving color correction across camera and illumination changes by contextual sample selection, *Journal of Electronic Imaging*, 21(2): 023015, 2012

[20] A. Alsam, J.Y. Hardeberg, Convex reduction of calibration charts. In Color Imaging X, Processing, Hardcopy, and Applications, SPIE Proceedings 5667, 38–46, 2005

Author Biography

Jon Y. Hardeberg received his MSc degree in signal processing from NTH (now NTNU) in Trondheim, Norway and his PhD from Ecole Nationale Supérieure des Télécommunications in Paris, France. He is currently Professor of Colour Imaging at Gjøvik University College,in Gjøvik, Norway, and member of the Norwegian Colour and Visual Computing Laboratory.

Margarita Khokhlova received her MSc degree in 3D Multimedia Technology in a collaborative master programme offered by Jean Monnet University in Saint-Etienne, France, and Gjøvik University College in Gjøvik, Norway. Currently she is a PhD student at the Université de Bourgogne in Dijon, France.