Pixel-wise illuminant estimation for mixed illuminant scenes based on near-infrared camera information

Joschua Thomas Simon-Liedtke, Jon Yngve Hardeberg, Per Ove Husøy; The Norwegian Colour and Visual Computing Laboratory, Gjøvik University College, Norway

Abstract

Computational color constancy or white balancing methods for digital cameras emulate the ability of the human visual system to adapt to different lighting situations and to maintain color constancy. Global white balancing algorithms have been shown to give remarkable results for scenes illuminated by one light source, but proven less adequate for multi-illumination scenes where multiple light sources are present. Using information from an additional near-infrared channel can be used to estimate the white point at every pixel in the image by comparing the pixels' NRGB values to a multi-dimensional lookup table with precomputed NRGB values. This estimated white point can then be used for white balancing via linearized Bradford transform. The lookup table requires measurement of multiple reflectance and illumination spectra that are representative for an office environment. The method performs better than conventional global white balancing methods.

Introduction

The ability of the human visual system (HVS) to adapt to different lighting situations such that spectrally flat reflectances appear neutral or "gray", is often referred to as color constancy [1].

Purves and Lotto explain color constancy as strategy of the HVS to cope with the inverse problem of retinal images based on an empirical process, in which "retinal images are linked to successful behavior by trial and error interactions with the environment" [2].

Digital cameras on the other hand use so-called color constancy methods that convert images taken under an unknown lighting condition into images taken under a known standard canonical light, usually D65 [3].

Global white balancing (WB) methods estimate a global white point (WP) and apply the same correction to all image points. The WP estimation can be performed in one of many existing color spaces, for example camera-RGB, XYZ or LMS color space [1, 3, 4]. Usually a diagonal 3×3 matrix is computed (f.ex. a Bradform transform [4]), but it is also possible to apply different transformations in different areas of the color space in the form of lookup tables (LUTs).

For multi-illuminant scenes, global white balancing methods often fail, and other image improvement methods like ACE [5], STRESS [6] or Retinex [7] and spatial white balancing methods have been proposed [8–12] including most recent research using a multi-illuminant random field (MIRF) [8] or empirical obtained spatial and spectral derivatives that can be used in a Lambertian image formation model [13]. In most spatial white balancing methods, the image is generally divided into smaller segments, on which global white balancing method are used.

The task of obtaining a good spatially dependent white point estimate is still a critical factor in contemporary imaging systems. Contemporary illuminants in office settings are typically designed to make objects appear to have the same color as would be seen with a natural daylight illumination. Due to energy efficiency requirements, however, the spectral emission of these illuminants is typically limited to the visible wavelength range. There is thus very often a large difference in emitted power in the near-infrared range when comparing artificial with natural light sources. The sun as well as old incandescent lamp types have a large amount of the emitted power in the near-infrared, while fluorescent lights and LEDs hardly have any.

This effect may help us in separating the mixture of the two types of illuminants in mixed illumination scenarios, in which mostly fluorescent and daylight light sources are present, which is very common for office settings. There, fluorescent light has become one of the most dominant light sources together with daylight from the outside, whereas incandescent light has began to disappear from office settings. Since this kind of settings are widely spread and important in the industry and other commercial working places, we decided on focusing on illuminant estimation in those kind of office settings.

As discussed above, the main difference of the spectra between fluorescent light and natural daylight can be seen in the emission of radiation in the near-infrared. Consequently, we assume that adding a fourth near-infrared channel will increase the performance of illuminant estimation in office settings. Fredembach *et al.* [14] presented an approach to this end that facilitates white point estimation by including near-infrared data from the scene, and comparing the average color values in sectors of the image to a database of previously measured light sources. A similar method will be presented in the following paper. For our method, we assume a scenario, where two main sources a present in the scene, a fluorescent and daylight light source. These two light sources are mixed together randomly in every image point. Consequently, the main task is how to separate the two illuminants in every pixel of the image.

General Methodology

We propose a pixel-wise white point estimation algorithm using a lookup table (LUT). The LUT is generated from empirical data recorded in the form of material spectral reflectances, illumination spectra and camera sensitivity curves. The LUT includes an additional near-infrared channel next to the traditional red, green and blue response data.

We propose an algorithm that retrieves the white point at every single pixel of the image in order to compute a diagonal white balancing matrix for each pixel individually. Likewise, we need the XYZ-values of the actual illuminant at each pixel and the canonical illuminant. It can be shown that these XYZ-values can be computed from the camera data represented in a vector that contains a value for near-infrared, red, green and blue (NRGB) with the help of a NRGB to XYZ calibration matrix determined by the camera manufacturer.

Mathematically, these NRGB values depend on the ray of light reaching the sensor from the each point in the scene and can be calculated by the following formula:

$$\begin{aligned} N(x,y) &= \int_\lambda C(x,y,\lambda)Q_N(x,y,\lambda)\mathrm{d}\lambda \\ R(x,y) &= \int_\lambda C(x,y,\lambda)Q_R(x,y,\lambda)\mathrm{d}\lambda \\ G(x,y) &= \int_\lambda C(x,y,\lambda)Q_G(x,y,\lambda)\mathrm{d}\lambda \\ B(x,y) &= \int_\lambda C(x,y,\lambda)Q_B(x,y,\lambda)\mathrm{d}\lambda \end{aligned} \tag{1}$$

In the formula, $C(x,y,\lambda) = \int_\lambda I(x,y,\lambda)R(x,y,\lambda)\mathrm{d}\lambda$ represents the color stimulus at each pixel, where $R(x,y,\lambda)$ are the reflectance components at each point, and $I(x,y,\lambda)$ represents the spectral distribution of the combination of the light sources. $Q_N(\lambda)$, $Q_R(\lambda), Q_G(\lambda)$, $Q_B(\lambda)$ are the camera sensor sensitivities for the red green blue and near-infrared channels.

We assume two main illuminants in the scene, one fluorescent and one daylight, of which the individual white points, WP_1 and WP_2, are known - for example after finding the maximum valued pixels in the image or by any other methods as proposed by Barnard *et al.* [1]. The illuminant at each pixel can be computed from the mixture combination coefficients (k_1,k_2) as expressed in the formula $I(x,y,\lambda) = k_1(x,y)\cdot I_1(\lambda)+k_2(x,y)\cdot I_2(\lambda)$, where $I_1(\lambda)$ indicates the SPD of the current daylight spectrum, $I_2(\lambda)$ indicates the SPD of the current fluorescent light spectrum. Both variables are measured in lux. In other words how much intensity of either one of the light sources are mixed together at each pixel in the image. $(k_1,k_2) = (0.2,0.3)$ means that a particular pixel gets 20% of the full intensity of light source 1 and 30% of the full intensity of light source 2. It can be shown that the normalized mixture combination coefficient $\alpha = \frac{k_1}{k_1+k_2}$ is enough to estimate the correct white point with normalized intensity at a given pixel (cf. Equation 2). This white point will then be used to compute the diagonal matrix for translating the pixel to canonical D65 lighting.

$$\mathsf{WP}^*_{\mathsf{mix}}(x,y) = \alpha(x,y)\cdot\mathsf{WP}_1 + (1-\alpha(x,y))\cdot\mathsf{WP}_2 \tag{2}$$

Thus the main steps of the proposed method can be summarized as following and visualized as in Figure 1. Firstly, the NRGB values of the two light sources are determined. Secondly, the NRGB of the image from the camera are converted into $\mathsf{XYZ}_{\mathsf{mix}}$ under the mixed illuminant using the camera specific calibration matrix. Thirdly, for each pixel the method evaluates the respective NRGB values and estimates the normalized mixture combination coefficient $\alpha = \frac{k_1}{k_1+k_2}$. Fourthly, this α is used to compute the normalized white point for the respective pixel (cf. Equation 2). Fifthly, a color adaptation transform (CAT) matrix is computed using the normalized white point and the linearized Bradford transform as proposed by Finlayson and Süsstrunk [4]. Eventually, the CAT matrix for each pixel is used to correct the $\mathsf{XYZ}_{\mathsf{mix}}$ co-ordinates under the the mixed illuminant into corrected $\mathsf{XYZ}_{\mathsf{D65}}$ under the standardized D65 lighting.

For more details on the proposed method, please read the master thesis by Simon-Liedtke [15].

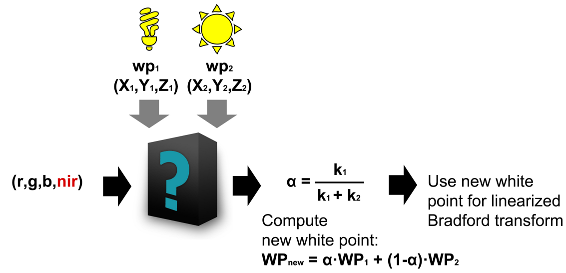

Figure 1. *The workflow of our proposed method visualized: The NRGB values at every pixel (x,y) are taken as input plus the XYZ co-ordinates of the original light sources present in the scene in order to compute the normalized mixture combination coefficients $\alpha = \frac{k_1}{k_1+k_2}$ for each pixel in the image. This will be used to compute the adapted new white point for every pixel, which serves as starting point for one of the conventional white balancing algorithms like the Bradford transform as in our implementation. The black box represents the most essential part of the workflow, namely the act of finding the most accurate mixture coefficient, i.e. the local white point estimation.*

Local White Point Estimation

In this section, we present the method that will estimate the normalized mixture combination coefficient α of the illumination from a multi-dimensional lookup table with precomputed NRGB values, precomputed from a representative set of object reflectance and light source spectra, in order to estimate the white point of the illumination at every pixel. Representative means that the lookup table spans the range of all possible NRGB values that can be obtained in a office setting with the respective normalized mixture combination coefficient as good as possible. Thus, the method can be divided into three parts: First, the collection of training data. Secondly, the building of the lookup table with the optimal NRGB values and normalized mixture combination coefficient α based on the collected training data under different illuminants and intensities. And thirdly, the retrieval of values from the lookup table.

Collection of training data

At first, a set of data is needed to precompute the NRGB values: The spectral reflectance data of a representative sets of office materials, a spectral power distribution data base of a representative set of office light sources and the camera sensitivity curves for each channel red, green, blue and near-infrared of the camera that is used.

Building the lookup table

Once the data sets are defined, NRGB values are precomputed for each material under different lighting situations for each pair of daylight-fluorescent light combination.

for all pairs of illuminant I_1 and I_2 in the database **do**
  for all reflectance spectra R_n in the database **do**
    for $k_1 = 0 : \varepsilon : 1.0$ **do**
      for $k_2 = 0 : \varepsilon : 1.0$ **do**
        $I_{mix}(\lambda) = k_1 \cdot I_1(\lambda) + k_2 \cdot I_2(\lambda)$
        $(N,R,G,B) = computeNRGB(Q_N(\lambda),\ldots$

$\dots Q_R(\lambda), Q_G(\lambda), Q_B(\lambda), R_n(\lambda), I_{mix}(\lambda))$ [1]

$\alpha = \frac{k_1}{k_1+k_2}$

$saveInCodeBook((N,R,G,B,\alpha))$

end for

end for

end for

end for

ε indicates how dense the entries will be in the lookup table. By default this value is set to 0.1.

Retrieval from the lookup table.

Once the lookup table has been created, it can be used within the white balancing workflow to retrieve the information for the correct α by finding the closest NRGB entry in the lookup table for every NRGB value in the image. If the lookup table does not have the exact NRGB value, the correct alpha can be obtained by different methods like for example nearest neighbor, averaging over k-nearest neighbors, interpolation etc.

Implementation and office reflectance data base

For the evaluation, we collected a set of spectral reflectances for representative office materials, light sources and a typical camera.

For the office reflectance database, we collected reflectance spectra of multiple materials that can be found in a normal office setting in the range of $400nm - 1000nm$ with an AvaSpec-ULS2048. We have a total of 1639 samples belonging to one of five main categories:

- Paints, varnishes & alloys: 1131
- Cloths: 44
- Human skin: 216
- Organic matter: 58
- Wild card / random samples: 190

For the light sources, we measured the spectra of five different fluorescent light sources and six daylight spectra, i.e. direct sunlight and indirect sunlight, with the very same spectrometer from Avantes.

Eventually, we measured the camera sensitivity for a Basler acA1300-30gm $\frac{1}{3}inch$ monochromatic progressive scan CCD sensor with the help of a Bentham monochromator TMc300. To simulate the channels, the overall sensitivity has been multiplied with the transmittances of blue, green, red and and near-infrared band-pass filters as provided by Midwest Optical Systems Inc. with peaks at $470nm$, $525nm$, $635nm$ and $785nm$.

In the implementation of the code book, we have a total of 7,622,637 entries computed from 5,727 reflectance measurements (including double, triple, etc. measurements) and 11 light sources.

Testing simulation and discussion

We created a set of 50 test images with an image size of 25×25 pixel. For each image, one fluorescent light source and one daylight source is chosen by random. For each pixel, one material is chosen randomly from the office reflectance database and a combination coefficient has been chosen depending on its (x,y) co-ordinates simulating illumination from the first gradient from the upper left corner and illumination from the lower right corner of the image.

The coefficient for the fluorescent light, k_1, depends on the distance from the top edge, and the coefficient for the daylight, k_2, depends on the distance from the right edge according to the following formula:

$$k_1(x,y) = \frac{Y-y}{2 \cdot Y}, \qquad k_2(x,y) = \frac{x}{2 \cdot X} \tag{3}$$

where Y and X are the dimensions of the image in the x- and y-axis.

The NRGB values have been computed using the mixture combination coefficients and the spectra of the material at each pixel with the help of the previously discussed Equation 1.

The normalized combination coefficients $\alpha(x,y)$ have been estimated finding the nearest neighbor entry to every NRGB pixel value in the lookup table and subsequently been used to correct the color of each pixel using the linear Bradford transform. Normalized means that the alpha lies between zero and one, reflecting only the ratio between the fluorescent and the daylight light source. If the α was not normalized, it would reflect the real intensity difference between the two light sources, which is unknown in this scenario.

Then, the difference between the theoretical reference XYZ_{D65} co-ordinates has been calculated according to the CIEDE2000 color difference formula [16] and compared to the color differences before correction, and color differences after correction using the average of both white points for a global linear Bradford transform, as can be seen in Figure 2. One random example on how the colors look before and after correction can be seen in the top row of Figure 2. The box plots in the bottom row below the different versions show the distribution of color differences between the reference (D65) and the different versions. The boxes show the location of 50% of CIEDE2000, and the "arms" show the location of 95% of the CIEDE2000s. It is therefore a good visualization of the variance of the distribution. The red dots represent the outliers of the CIEDE2000.

Comparing the results of the color references reveals that our proposed method has a lower average color difference than any other tested method, and that 95% of the the samples have a color difference within or lower than the average than without correction. Generally, it can be seen that the multi-dimensional lookup has a lower average and a lower spreading of errors than the original or the global correction method. Moreover, improvement can be seen visually between the different versions (cf. Figure 2.c.)). Namely, we can observe somewhat neutral colors in our proposed method, whereas both the uncorrected version and the corrected version using the global average WP method have visible color tints.

The method could be improved by using other methods than k-nearest neighbour to retrieve the values from the lookup table like for example interpolation. "Cleaning up" the lookup table might furthermore improve the performance and reduce the number of outliers. Also, we consider using the method not pixel-wise

[1] see Equation 1

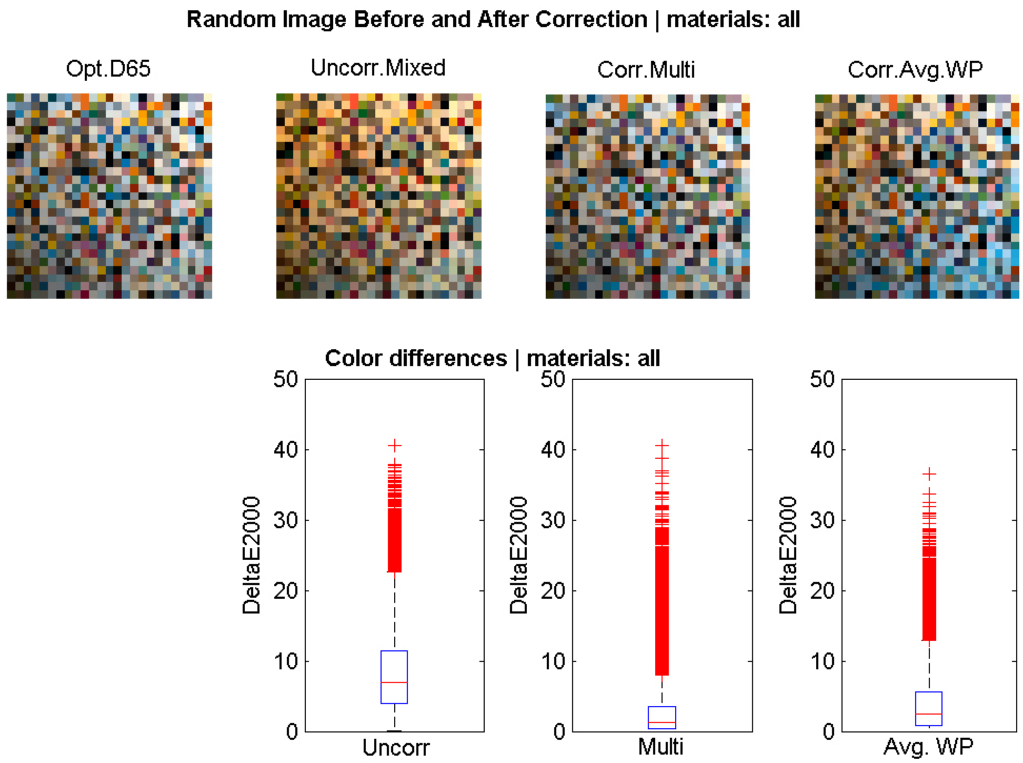

Figure 2. *The images show a virtual test image using spectra from real live materials under simulated fluorescent lighting from the the top left and daylight lighting from the bottom right. The first image, a.), serves as a reference as how the image should look like after perfect white balancing and confirmation to D65. The second image, b.), shows the uncorrected version of the image with uncorrected condition under mixed illumination. The third image, c.), shows the corrected image using our proposed method based on multi dimensional lookup method. And the fourth image, d.) shows the corrected image using a global white balancing method based on a average white point. It can be seen that our proposed method is visually closest to the reference image. Moreover, the distribution of the color difference to the reference shows that it has the lowest average CIEDE2000 and the lowest standard deviation of all tested methods.*

but rather on segments of the image in order to reduce computational costs.

It mind be interesting to test whether or not the near-infrared data is necessary at all or if similar results could be obtained when only the RGB information is at hand. And more importantly, the method should be verified on live action images as well.

Conclusion and Perspectives

We proposed a workflow of pixel-wise white balancing based on the linearized Bradford transform. For the estimation of the white points, a lookup table is built with the precomputed camera response values obtained from databases, containing reflectance spectra data from typical office materials, the spectral power distributions of typical office light sources, and the measured channel sensitivities of a four channel camera with sensitivities within the red, green, blue and near-infrared. The values are precomputed for different combinations of the light sources and the exact mixture combination is stored within the lookup table. For each pixel of the image, the respective normalized mixture combination coefficient can be obtained by finding the closest NRGB entry in the database and using it to compute the white point at the respective pixel assuming that the white point of the main light sources are known.

In future work, it will be important to test the algorithm on live action images, and compare it to other white balancing methods. Also, using other evaluation methods for assessing color image quality like psychometric experiments should be considered. Methods to obtain NRGB data have been proposed for example by Fredembach and Süsstrunk [17, 18].

Acknowledgements

We especially want to thank Cisco Systems Norway AS for their support in this project and the Erasmus Mundus Master programme Color in Informatics and MEdia Technology (CIMET) financed by the European Union (EU). We also say thanks for all the participants at the Gjøvik University College who let us measure the reflectance of their skin.

References

[1] Kobus Barnard, Lindsay Martin, Adam Coath, and Brian Funt, "A Comparison of Computational Color Constancy Algorithms - Part II: Experiments with Image Data," *IEEE Transactions on Image Processing*, vol. 11, pp. 985–995, 2002.

[2] Dale Purves and R. Beau Lotto, *Why We See What We Do Redux: A Wholly Empirical Theory of Vision*, Sinauer Associates, Inc., 2011.

[3] Kobus Barnard, Lindsay Martin, Adam Coath, and Brian

Funt, “A Comparison of Computational Color Constancy Algorithms - Part I: Methodology and Experiments With Synthesized Data,” *IEEE Transactions on Image Processing*, vol. 11, pp. 972–983, 2002.
[4] Graham D. Finlayson and Sabine Süsstrunk, “Spectral Sharpening and the Bradford Transform,” in *Proc. Color Imaging Symposium (CIS 2000)*, Derby, UK, 2000, University of Derby, pp. 236–243.
[5] Alessandro Rizzi, Carlo Gatta, and Daniele Marini, “A new algorithm for unsupervised global and local color correction,” *Pattern Recogn. Lett.*, vol. 24, no. 11, pp. 1663–1677, July 2003.
[6] Øyvind Kolås, Ivar Farup, and Alessandro Rizzi, “Spatio-Temporal Retinex-Inspired Envelope with Stochastic Sampling: A Framework for Spatial Color Algorithms,” *Journal of Imaging Science and Technology*, vol. 55, no. 4, pp. 40503–1–40503–10, 2011.
[7] Edwin H. Land and John J. McCann, “Lightness and retinex theory,” *Journal of the Optical Society of America*, vol. 61, no. 1, pp. 1–11, 1971.
[8] Shida Beigpour, Christian Riess, Joost van de Weijer, and Elli Angelopoulou, “Multi-Illuminant Estimation with Conditional Random Fields,” *IEEE Transactions on Image Processing*, vol. 23, pp. 83–96, 2014.
[9] Ching-Chih Weng, H. Chen, and Chiou-Shann Fuh, “A novel automatic white balance method for digital still cameras,” in *IEEE International Symposium on Circuits and Systems*, Kobe, Japan, 2005, IEEE, pp. 3801–3804.
[10] Arjan Gijsenij, Rui Lu, and Theo Gevers, “Color Constancy for Multiple Light Sources,” *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 697–707, Feb. 2012.
[11] Feng-Ju Chang, Soo-Chang Pei, and Wei-Lun Chao, “Color constancy by chromaticity neutralization,” *Journal of the Optical Society of America A, Optics and Image Science*, vol. 29, no. 10, pp. 2217–2225, Oct 2012.
[12] Ivaylo Boyadzhiev, Kavita Bala, Sylvain Paris, and Frédo Durand, “User-guided white balance for mixed lighting conditions,” *ACM Trans. Graph.*, vol. 31, no. 6, pp. 200:1–200:10, Nov. 2012.
[13] Jan-Mark Geusebroek, Rein van den Boomgaard, Arnold WM Smeulders, and Theo Gevers, “Color constancy from physical principles,” *Pattern Recognition Letters*, vol. 24, no. 11, pp. 1653–1662, 2003.
[14] Clément Fredembach and Sabine Süsstrunk, “Illuminant estimation and detection using near-infrared,” in *Proceedings of SPIE Electronic Imaging: Digital Photography V*, 2009, vol. 7250 of *Proc. of Electronic Imaging*.
[15] Thomas Joschua Simon-Liedtke, “Using Near-Infrared for Color Correction of Multi-Illumination Scenes in Video Conferencing Systems,” M.S. thesis, Gjøvik University College, 2013.
[16] Gaurav Sharma, Wencheng Wu, and Edul N. Dalal, “The CIEDE2000 Color-Difference Formula: Implementation Notes, Supplementary Test Data, and Mathematical Observations,” *Color Research & Application*, vol. 30, no. 1, pp. 21–30, 2005.
[17] Clément Fredembach, Nathalie Barbuscia, and Sabine Süsstrunk, “Combining visible and near-infrared images for realistic skin smoothing,” in *Proc. IS&T/SID 17th Color Imaging Conference (CIC)*, 2009, pp. 242–247.
[18] Sabine Süsstrunk and Clément Fredembach, “Enhancing the Visible with the Invisible: Exploiting Near-Infrared to Advance Computational Photography and Computer Vision,” in *SID Symposium Digest of Technical Papers*. 2010, vol. 41 of *1*, pp. 90–93, SID.

Author Biography

Joschua Simon-Liedtke did his understudies in Audiovisual Media in Stuttgart, before he received a Master of Science in Colour in Informatics and MEdia Technology (CIMET), which he attended in Granada, Saint-Etienne and Gjøvik in 2013. He is currently working as Ph.D.-candidate in the field of Color Image Quality Improvement for Color Deficient People at the Gjøvik University College, where he is a member of the Norwegian Colour and Visual Computing Laboratory. The main focus of his research work is Color Image Enhancement Algorithms and the Human Visual System.

Jon Y. Hardeberg received his Ph.D. from the ENST in Paris in 1999. After a short industry career in the US, he joined Gjøvik University College in Norway in 2001, where he is Professor of Colour Imaging and member of the Norwegian Colour and Visual Computing Laboratory. His broad research interest within the field has resulted in more than 150 papers. He currently leads several large research projects, including the Initial Training Network CP7.0.

Per Ove Husøy received his Ph.D. from NTH, Norway in 1992. He has subsequently worked with signal processing and pattern recognition technology for several companies and application areas: radar, audio, video and image. He currently works as acoustic consultant at Brekke & Strand acoustics.