Estimation of Reflectance Spectra Using Multiple Illuminations

Ville Heikkinen, Tuija Jetsu, Jussi Parkkinen and Timo Jääskeläinen, University of Joensuu (Finland); and Reiner Lenz, Linköping University (Sweden)

Abstract

Using ordinary digital cameras as relatively cheap measurement devices for estimating spectral color properties has become an interesting alternative to making pointwise high precision spectral measurements with special equipments like photospectrometers. The results obtained with these methods cannot compete with the quality of the traditional high resolution devices but they are very attractive since the equipment is relatively cheap and instant measurements are obtained for millions of measurement points.

In this paper we investigate the problem of estimating reflectance spectra from measurements taken with ordinary digital RGB cameras. We study the effects of using multiple illuminations and treat the estimation of the reflectance spectra as a regression or a statistical inversion problem. We use both, linear- and non-linear estimation methods where we focus on using reproducing kernels to avoid explicit formulation of non-linearities. We also include non-linear conditions based on the properties of the reflection spectra. Munsell Matte color and Pantone are used as data sets to support the proposed methods. The experiments show that the proposed methods improve the estimation results when compared to standard linear methods.

Introduction

This study considers the estimation of reflectance spectra from the responses of color recording device. Reflectance information is independent of device and illumination and therefore it can be beneficial for several applications. Some examples are color calibration of devices and quality control applications. It has been also used in art painting recording [1]. Typical devices used in these applications are digital cameras and color scanners.

The estimation models for digital camera or scanner data can be constructed using the same methods. The properties of incoming spectra are usually more varied for the digital camera. For the digital camera, there is also an additional problem of largely varying recording illumination. Possibility to measure the device responses of object under different illuminations can be also used as an asset to increase the accuracy of estimations [3].

In this paper we continue to investigate the problem of estimating reflectance spectra from multiple illumination measurements taken with a digital RGB-camera. We use both, linear- and non-linear estimation methods and introduce non-linear transformation based on the properties of the reflectance spectra.

Estimation methods

Digital color recording devices capture the spectrum of physical stimuli by filtering the incoming color signal through color filters with different spectral transmittances. In the following we denote matrices as capital letters and vectors as boldface letters. Camera model for sensor responses can be approximated as

$$\mathbf{x} = W\mathbf{r}, \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^k$ is the response vector, $W \in \mathbb{R}^{k\times n}$ is the response matrix and $\mathbf{r} \in [0,1]^n$ is the reflectance. The response matrix W depends on both, illumination and spectral sensitivities. In this study we use an ordinary RGB-camera with one illumination source ($k = 3$) or two illumination sources ($k = 6$). The goal of this study is to find inversion operator $L : \mathbb{R}^k \to [0,1]^n : \mathbf{x} \to \hat{\mathbf{r}}$ which provides an estimation of the sampled reflectance. We treat the estimation of reflectance spectra $\mathbf{r}$ as a regression problem or an inverse problem.

The standard methods used for estimation are linear models incorporating a priori knowledge [1],[3]. In the case of additive normally distributed noise the statistical inversion problem formulation leads to the well-known Wiener estimation

$$\hat{\mathbf{r}}_w = \Sigma_{rr} W^T (W\Sigma_{rr} W^T + \gamma I_k)^{-1}\mathbf{x} \tag{2}$$

where camera sensitivities, illumination and the correlation matrix Σ_{rr} of training ensemble are used [6]. The parameter $\gamma \geq 0$ depends on the identically and independently distributed noise variance in responses.

If we have a priori information of the reflectance spectra and the corresponding camera responses we can perform the estimation with linear or non-linear regression. In this case explicit knowledge of the response matrix is not needed. It has been shown that non-linear regression can lead to increased accuracy of estimates when the dimension of the subspace of reflectance is higher than the number of response functions of the observer (e.g. see [2],[4],[5]).

For non-linear regression we use methods from the theory of Reproducing Kernel Hilbert Spaces to avoid the explicit formulation of the non-linear mapping [4], [7]. This can be more practical than the explicit usage of higher order polynomial terms of response vector $\mathbf{x} = (x_1, \ldots, x_k)^T$ as

$$\Phi_p(\mathbf{x}) = (x_1, x_2, \ldots, x_k, x_1^2, x_2^2, \ldots, x_k^2, \ldots)^T$$

that are only practical in lower dimensional response spaces [2], [5]. In the case of kernels the explicit non-linear feature mapping $\Phi : \mathbb{R}^k \to \mathbb{R}^N$ of the possibly high-dimensional response vector $\mathbf{x}$ is avoided by using a chosen kernel function κ. Dimension N of the feature space depends on the kernel and can be infinite. Estimate of reflectance corresponding to $\mathbf{x}$ is formulated in this study using linear functions in the feature space. The function κ defines an inner product in the feature space $\kappa(\mathbf{x},\mathbf{z}) = \langle \Phi(\mathbf{x}), \Phi(\mathbf{z}) \rangle$ [7]. If Kernel Ridge Regression algorithm is used separately for all the n spectral samples, the estimate computed from $\mathbf{x}$ is a linear combination of all training spectra [4]. These separate scalar valued regression problems correspond to minimization problem for matrix $F \in \mathbb{R}^{n\times N}$:

$$\arg\min_F \left(\sum_i^l \|\mathbf{r}_i - F\Phi(\mathbf{x}_i)\|_2^2 + \gamma \|F\|^2 \right), \tag{3}$$

where l is the size of training set and parameter $\gamma \geq 0$ allows to trade-off between the norm of the solution and the datafitting

accuracy. Using kernel function $\kappa(\mathbf{x},\mathbf{x}_i) = \langle \Phi(\mathbf{x}), \Phi(\mathbf{x}_i) \rangle$, the feature space is used only implicitly and estimate corresponding to response $\mathbf{x}$ becomes

$$\hat{\mathbf{r}} = R^T (K + \gamma I)^{-1} \kappa^{\mathbf{x}}. \quad (4)$$

where $K_{jm} = \kappa(\mathbf{x}_j, \mathbf{x}_m)$, $\kappa^{\mathbf{x}} = (\kappa(\mathbf{x},\mathbf{x}_1), \ldots, \kappa(\mathbf{x},\mathbf{x}_l))^T$ and $R \in \mathbb{R}^{l \times n}$ is matrix of training spectra [4].

The solution of the regression problem (4) can be seen to be related also to the inverse problem solution (2). The Wiener estimate (2) with $\gamma > 0$ can be written in the form

$$\begin{aligned} \hat{\mathbf{r}}_w &= \Sigma_{rr} W^T (W \Sigma_{rr} W^T + \gamma I_k)^{-1} \mathbf{x} \\ &= R^T R W^T (W R^T R W^T + \gamma I_k)^{-1} \mathbf{x} \\ &= R^T (R W^T W R^T + \gamma I_l)^{-1} R W^T \mathbf{x} \\ &= R^T (K + \gamma I_l)^{-1} \kappa_w^{\mathbf{x}}, \end{aligned} \quad (5)$$

where the third equality can be obtained using SVD (Singular Value Decomposition [8]) of the matrix RW^T. The linear kernel is now defined in spectral space and we have for the elements of matrix K:

$$K_{jm} = \kappa_w(\mathbf{r}_j, \mathbf{r}_m) = \langle W\mathbf{r}_j, W\mathbf{r}_m \rangle = \mathbf{r}_j^T W^T W \mathbf{r}_m$$

and

$$\kappa_w^{\mathbf{x}} = ((W\mathbf{r}_1)^T \mathbf{x}, \ldots, (W\mathbf{r}_l)^T \mathbf{x})^T.$$

The kernel κ_w can be replaced by non-linear versions leading to applications of a higher dimensional feature space. Some examples of reproducing kernels are: polynomial kernels of different degrees, splines, Gaussian, etc. [7]. For this study we chose the Gaussian kernel with

$$K_{jm} = \kappa_g(\mathbf{r}_j, \mathbf{r}_m) = \exp\left(-\frac{\|W\mathbf{r}_j - W\mathbf{r}_m\|_2^2}{2\sigma^2} \right),$$

and

$$(\kappa_g^{\mathbf{x}})_j = \exp\left(-\frac{\|W\mathbf{r}_j - \mathbf{x}\|_2^2}{2\sigma^2} \right),$$

where $\sigma^2 > 0$ is a free parameter that enables us to control the model.

In our approach we use also a non-linear coordinate transformation to include the a priori knowledge that results are reflectance spectra with function values between zero and one. We use the following non-linear transformation for reflectance spectra

$$\tilde{\mathbf{r}} = \operatorname{arctanh}(2\mathbf{r} - 1), \quad (6)$$

where arctanh : $[-1,1] \to R$ is used pointwise. Restoration from non-linear features is then given by

$$\mathbf{r} = (1 + \tanh(\hat{\mathbf{r}}))/2. \quad (7)$$

In this way the reflectance estimates are constrained in the region $[0,1]$, because tanh : $R \to [-1,1]$. Actual response is the result of a linear process, if we assume that the system (1) is free of non-linearities. The application of the transformation (6) introduces however a non-linearity into the whole estimation process. The atanh-transformation can be used for Gaussian kernel as

$$\hat{\mathbf{r}} = \tilde{R}^T (K + \gamma I_l)^{-1} \kappa^{\mathbf{x}}, \quad (8)$$

and for Wiener estimation as

$$\hat{\mathbf{r}} = \tilde{R}^T R W^T (W \Sigma_{rr} W^T + \gamma I_k)^{-1} \mathbf{x}, \quad (9)$$

where $\tilde{R}$ denotes atanh-transformed training spectra as rows.

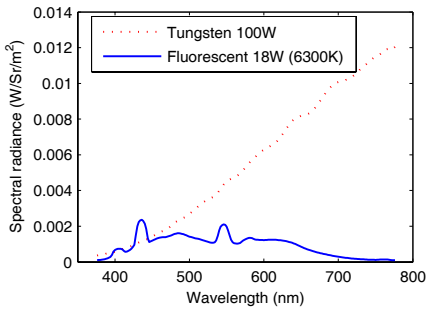

Figure 1. *Fluorescent and tungsten illuminations.*

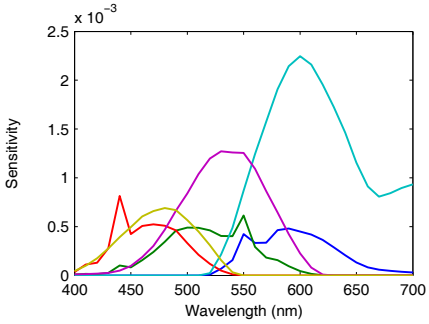

Figure 2. *Sensitivities under fluorescent and tungsten illuminations.*

Experiments

We chose to use the Munsell Book of Color - Matte Finish Collection (1269 samples) and Pantone (922 samples) as data sets in experiments. We validate our results on two different test cases. In the first case responses correspond to camera responses simulated under two different illumination conditions: Philips DeLuxe90 TLD 18W/965 6300K lamp and tungsten 100W lamp sources (Figure 1). Estimation with multiple illuminations has been also performed using led-illuminations [3]. All spectral distributions were sampled from 400 nm to 700 nm with 10 nm steps. Normal distributed noise was added to test responses corresponding to 43 dB SNR. The Wiener method (2)without atanh-transformation or noise (SNR $= \infty$) in the test responses was used as a reference method in this study. This provides an upper bound for the performance of the linear Wiener system under ideal simulated conditions and corresponds to an estimate (2) with $\gamma = 0$.

We use the estimated spectral sensitivities of a digital consumer camera [6]. The spectral sensitivities under two illuminations are depicted in Figure 2. Test responses and the known response functions $\{w_i\}_{i=1}^k$ were scaled with a diagonal matrix S, where

$$S_{ii} = \frac{1}{\|w_i\|}, \; i = 1, \ldots, k.$$

The kernel is given by

$$\kappa_g(\mathbf{r_j}, \mathbf{r_m}) = \exp\left(-\frac{1}{2\sigma^2} \|SW\mathbf{r_j} - SW\mathbf{r_m}\|_2^2 \right).$$

In the second setting we used actually measured RGB responses using the same tungsten and fluorescent illumination sources. The objects measured were the chips from the Munsell Book of Color - Matte Finish Collection. We captured JPEG images of the Munsell chips with a Fujifilm Finepix S1 Pro digital camera. We used a Nikon AF Nikkor 25-50 mm zoom lens. The spectral sensitivities of this camera are different from the ones used in the simulated case and unknown. Also the non-linearities

of the camera are not known. Only response values from the regions $[5,250]$ were used in the experiments. The kernel for measured setting is defined in terms of noisy training responses

$$\kappa_g(\mathbf{x_j},\mathbf{x_m}) = \exp\left(-\frac{1}{2\gamma^2}\|S\mathbf{x_j} - S\mathbf{x_m}\|_2^2\right),$$

where $S_{ii} = 1/255$. We compare the proposed non-linear regression method to the linear least squares regression. Linear pseudoinverse estimation is formally written as

$$\hat{\mathbf{r}} = R^T X(X^T X)^{-1}\mathbf{x}, \tag{10}$$

where $X \in \mathbb{R}^{l\times k}$ denotes matrix of response values from training set.

All results are evaluated using RMSE-error

$$\text{RMSE} = (\|\mathbf{r} - \hat{\mathbf{r}}\|^2/n)^{\frac{1}{2}}, \tag{11}$$

where $\|\cdot\|$ denotes the Euclidean norm and n is the length of the vectors $\mathbf{r}$ and $\hat{\mathbf{r}}$ corresponding to original and estimated spectrum. We use the following abbreviations in tables: Avg. = average error of the test set, Std. = standard deviation of the error, Max. = maximum error in the test set. The case where noise is present in the response values is denoted as (n) and when we use the atanh-transformation we write (t). The results for estimations are compared against PCA approximations of different ranks for the test set which are presented in Table 1.

In every case the free parameters σ and γ of the estimation model (8) with the Gaussian kernel were searched using an N-fold cross validation routine for the training set [8]. In N-fold cross validation, the training set is first separated into N equal-sized parts. The model is then trained using $N-1$ parts and tested with remaining portion. This procedure is repeated for all the N parts separately. We chose $N = 10$. Using certain parameter values, the average value of the chosen error value is calculated from all ten test set separations. When this method is used, also the response values corresponding to the training spectra are required. Parameter sets for the Gaussian kernel where chosen using intervals $\sigma \in [0.5,4]$ and $\gamma \in [10^{-3},10^{-2}]$. The parameter values corresponding to the minimum of average RMSE-error from cross-validation was chosen.

In the first case for the simulated setting, the training and test set separation was done using the two subsets: Munsell I (669 samples) & Munsell II (600 samples). The estimation results obtained for the simulated case for the Munsell I / Munsell II case are collected in Table 2. These results are averaged over 10 random partitions of the Munsell set into training (669 samples) and (separate) test sets (600 samples). It can be seen that without noise, without atanh-transformation and using only one illumination, the accuracy of the Gaussian kernel is comparable to 4-dimensional PCA approximation in terms of average RMSE. Error values are halved compared to the ideal Wiener estimate. Tungsten illumination leads to better performance than fluorescent illumination. When the noise and the atanh-transformation are used the error values increase but are still significantly lower compared to the ideal Wiener estimate. When both illuminations are used, the performance of the ideal Gaussian model shows again an improvement over the ideal Wiener result. When the noise and atanh-transformation are introduced for the Gaussian kernel, the performance of the model and the ideal Wiener model are equal. The Wiener model combined with the atanh-transformation lead to poor results for every illumination condition.

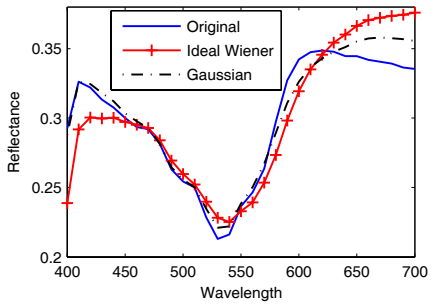

Figure 3. *An example of an estimate for a Munsell spectrum using Wiener and Gaussian kernel estimations. Training set: Munsell I, Illumination: Tungsten.*

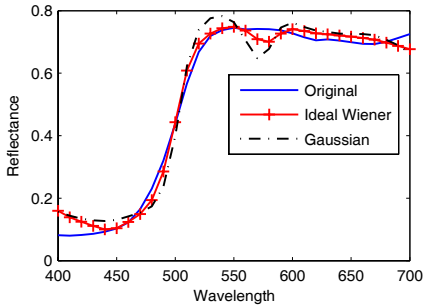

Figure 4. *An example of an estimate for a Munsell spectrum using Wiener and Gaussian kernel estimations. Training set: Pantone, Illumination: Tungsten & fluorescent.*

In second case for simulated setting the Pantone (922 samples) and the Munsell (1269 samples) were used as a training and test sets, respectively. The results for the Pantone / Munsell setting in Table 3 show significantly decreased performance from previous setting for both estimation models. Using one illumination, and the Gaussian kernel, model again improves from the ideal Wiener estimate in terms of average RMSE. Maximal error values are lower for the ideal Wiener estimate. The Gaussian model with atanh-transformation and noise gives similar performance as the ideal estimation. When both illuminations are used the ideal Wiener estimate leads to an improved performance over Gaussian model. The method used for the estimation of parameters has overfitted the training population and doesn't generalize well. If the intervals for the parameters were chosen to include even smaller values, the overfitting effect might be stronger. When the σ parameter is chosen to be larger, the performance of the Gaussian model approaches to the ideal Wiener model. In this sense the routine for searching the model parameters doesn't seem to be appropriate when the training and test populations are not similar. Again, the Wiener model combined with the atanh-transformation decreases the performance significantly from ideal model.

The estimation results obtained for the measured responses using regression estimations are collected in Table 4. These results are averaged over 10 random partitions of the Munsell set into training (669 samples) and (separate) test sets (495 samples). It can be seen that in this measured setting the linear pseudoinverse method without any constraints lead to poor estimations. This is probably due to the non-linearities in the system. When the atah-transformation is used for the linear pseudoinverse, the estimation performance increases significantly. This is opposite effect when compared to simulated setting where the system was free of non-linearities. For the Gaussian kernel the results are

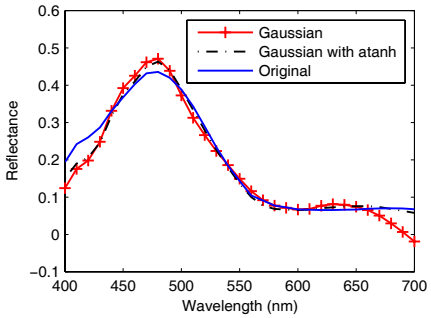

Figure 5. *An example of an estimate for a Munsell spectrum using Gaussian kernel with and without the atanh-transformation. Training set: Pantone, Illumination: Tungsten & fluorescent.*

similar with or without the atanh-transformation. Performance of the kernel improves from the 3-dimensional PCA-approximation and significantly improves from the pseudoinverse estimation.

Table 1. RMSE errors for Munsell sets using different PCA approximations. Results are average of ten randomizations of 600 samples.

Rank	Avg.	Std.	Max.
$k=3$	0.0190	0.0139	0.1066
$k=4$	0.0129	0.0097	0.0702
$k=5$	0.0092	0.0056	0.0453
$k=6$	0.0075	0.0042	0.0281
$k=7$	0.0055	0.0033	0.0285
$k=8$	0.0045	0.0025	0.0169

Table 2. Average RMSE errors for the simulation setting Munsell I/Munsell II (Average of 10 randomizations for noise and sets). Illumination sources: Fluorescent and/or Tungsten.

Method	Avg.	Std.	Max.
Fluorescent			
Ideal Wiener	0.0230	0.0186	0.1534
Ideal Wiener (t)	0.0392	0.0198	0.1745
Ideal kernel κ_g,	0.0119	0.0097	0.0745
Kernel κ_g (n, t)	0.0124	0.0103	0.0849
Tungsten			
Ideal Wiener	0.0211	0.0154	0.1063
Ideal Wiener (t)	0.0378	0.0157	0.1171
Ideal kernel κ_g,	0.0103	0.0084	0.0653
Kernel κ_g (n, t)	0.0111	0.0087	0.0715
Fluorescent & Tungsten			
Ideal Wiener	0.0090	0.0052	0.0422
Ideal Wiener (t)	0.0324	0.0130	0.0828
Ideal kernel κ_g,	0.0066	0.0048	0.0406
Kernel κ_g (n, t)	0.0091	0.0048	0.0397

Discussion

In this study we have presented results when linear Wiener estimation and linear least squares regression models are extended to non-linear using the Gaussian kernel. We have also presented physically motivated constraint of reflectance estimate to [0,1] using non-linear transformation in spectral space. Experimental results have been derived for a digital RGB camera using tungsten and fluorescent illuminations. The methods can also be used with general color scanner and also for arbitrary

Table 3. Average RMSE errors for the simulation setting Pantone/Munsell (Average of 10 randomizations for noise and sets). Illumination sources: Fluorescent and/or Tungsten.

Method	Avg.	Std.	Max.
Fluorescent			
Ideal Wiener	0.0357	0.0172	0.1487
Ideal Wiener (t)	0.0438	0.0175	0.1557
Ideal kernel κ_g,	0.0326	0.0209	0.1721
Kernel κ_g (n, t)	0.0333	0.0217	0.1759
Tungsten			
Ideal Wiener	0.0301	0.0145	0.1158
Ideal Wiener (t)	0.0396	0.0132	0.1088
Ideal kernel κ_g,	0.0273	0.0165	0.1186
Kernel κ_g (n, t)	0.0278	0.0167	0.1277
Fluorescent & Tungsten			
Ideal Wiener	0.0133	0.0090	0.0503
Ideal Wiener (t)	0.0299	0.0085	0.0614
Ideal kernel κ_g,	0.0184	0.0101	0.0692
Kernel κ_g (n, t)	0.0208	0.0099	0.0759

Table 4. Average RMSE errors for measured setting Munsell I/Munsell II setting (Average of 10 randomizations for sets). Illumination sources: Fluorescent and/or Tungsten.

Method	Avg.	Std.	Max.
Fluorescent			
Lin. pseudoinverse (n)	0.0554	0.0240	0.1669
Lin. pseudoinverse (n, t)	0.0263	0.0186	0.1154
Kernel κ_g (n)	0.0163	0.0115	0.0827
Kernel κ_g (n, t)	0.0161	0.0117	0.0854
Tungsten			
Lin. pseudoinverse, (n)	0.0547	0.0228	0.1629
Lin. pseudoinverse (n, t)	0.0270	0.0176	0.1145
Kernel κ_g (n)	0.0159	0.0108	0.0784
Kernel κ_g (n, t)	0.0163	0.0118	0.0837
Fluorescent & Tungsten			
Lin. pseudoinverse, (n)	0.0503	0.0242	0.1644
Lin. pseudoinverse (n, t)	0.0226	0.0160	0.1091
Kernel κ_g (n)	0.0140	0.0103	0.0768
Kernel κ_g (n, t)	0.0136	0.0102	0.0795

color transformations. It was found that the cross-validation routine for searching the model parameters wasn't appropriate when the training and test populations were different. Parameter values have to be constrained in order to avoid overfitting to training set. When multiple illuminations are used, the ideal Wiener estimation seem to provide accurate estimation for the chosen sets and sampling interval. It should be noted that efficient use of the Wiener method also requires the estimation of noise variance in real conditions. For linear models the atanh-transformation leading to physically feasible estimation cannot be used in general. In this study, use of the atanh-transformation improved the results for real measurements because of non-linearities in the camera system.

References

[1] H. Haneishi, T. Hasegawa, A. Hosoi, Y. Yokoyama, N. Tsumura and Y. Miyake, System design for accurately estimating the reflectance spectra of art paintings, Appl. Opt. **39**, 6621-6632 (2000).

[2] D. R. Connah, J. Y. Hardeberg, Spectral recovery using polynomial models, Proceedings of the SPIE Vol. 5667, pg. 65-75. (2005).

[3] J-I. Park, M-H. Lee, M. D. Grossberg and S. K. Nayar, Multispectral

Imaging Using Multiplexed Illumination, ICCV 2007.
[4] V. Heikkinen, T. Jetsu, J. Parkkinen, M. Hauta-Kasari, Timo Jaaskelainen and S.D. Lee, Regularized learning framework in estimation of reflectance spectra from camera responses, J. Opt. Soc. Am. A, Vol.24 No.9, (2007).
[5] T. Jetsu, V. Heikkinen, J. Parkkinen, M. Hauta-Kasari, B. Martinkauppi, S.D. Lee, H.W. Ok, and C.Y. Kim, Color calibration of digital camera using polynomial transformation, in CGIV, Third European Conference on Color in Graphics, Imaging and Vision, IS&T, USA, pg. 163-166, (2006).
[6] M. Solli, M. Andersson, R. Lenz, B. Kruse, Color Measurements with a Consumer Digital Camera Using Spectral Estimation Techniques, SCIA, LNCS 3540, pg. 110-114, (2005).
[7] B. Schölkopf and A.J. Smola, Learning With Kernels, MIT Press, Cambridge, 2002.
[8] T. Hastie, R. Tibshirani and J. Friedman, The Elements of Statistical Learning: Data mining, Inference and Prediction Springer-Verlag, New York, 2001.

Author Biography

Ville Heikkinen received his MSc. degree in Applied Mathematics from University of Joensuu, Finland in 2004. Currently he is a PhD. student and member of color research group at University of Joensuu.