Non-negative matrix factorization for spectral colors using genetic algorithms: Substantially Updated Version

Vladimir A. BOCHKO[†], Kimberly A. JAMESON[††], Toshiya NAKAGUCHI[†††], Yoichi MIYAKE[†††], *and* Jarmo T. ALANDER[†],

SUMMARY In this paper we introduce a novel method for non-negative matrix factorization (NMF) using a genetic algorithm. The method finds the optimal basis functions for the spectral colors in both spectral and color spaces. We show that one version of the proposed algorithm works as well as the standard NMF algorithm in spectral space. Further, this algorithm is modified to obtain a functionality to work in color space which the standard NMF is currently not capable of providing. The modification involves optimization in color space reducing the approximation error by a factor of 6 for Macbeth ColorChecker colors. The algorithm can be used in digital camera design. In addition, we propose an algorithm based on multiobjective optimization in both spectral and color space which can be used in digital image archiving.

key words: *Genetic algorithm, spectral colors, Macbeth ColorChecker, non-negative matrix factorization*

1. Introduction

Non-negative Matrix Factorization (NMF) gives a representation of the latent data structure and reduces the dimensionality of a dataset. There are at least two good physical reasons to use the NMF technique for spectrometry and imaging: First, optical sensors generate non-negative signals, that may benefit from non-negative decomposition, and, second, the optical color filters can be physically implemented only with non-negative spectral characteristics [1]. In multispectral imaging, for example, the technique may be used to calculate digital camera spectral band characteristics and is particularly attractive for spectral reflectance estimation using the Wiener method [2].

NMF solves the problem of non-negative data decomposition [3–6]. NMF has been used in spectral unmixing for non-resolved space object characterization [3]. NMF and a related technique called non-negative tensor factorization (NTF) were also used to define the optimal non-negative representation of spectral colors for different sets including the Macbeth ColorChecker (MCC) and Munsell color set [1, 7]. The present NMF methods represent measured spectra by a few non-negative basis functions. NMF determines non-negative factors $\boldsymbol{W}$ and $\boldsymbol{Z}$ and implements non-negative factorization of a given matrix $\boldsymbol{X}$ as follows: $\boldsymbol{X} \approx \boldsymbol{W}\boldsymbol{Z}$. The size of $\boldsymbol{X}$ is the number of wavelengths times the number of spectra, the columns of $\boldsymbol{W}$ are the basis functions found to minimizing the error of the approximation, and $\boldsymbol{Z}$ is a matrix of weights.

The present work investigates a common form of 3-primary filter systems that are standard in the digital camera industry. Practically speaking, if spectral response profiles of camera color sensors are designed to equal color matching functions of a human Standard Observer (such as $\bar{x}$, $\bar{y}$, $\bar{z}$ of the CIE 1931 XYZ color system), then fidelity between a digital image and a scene's original color should be achieved. In practice, however, such designs do not yield correct digital color capture and reproduction. One reason for such color fidelity failures is that display device controls commonly employ linearly transformed color matching functions $\bar{r}$, $\bar{g}$, $\bar{b}$ of the CIE 1931 RGB color system. Such transformations pose a problem because, while the positive portion of $\bar{r}$, $\bar{g}$, $\bar{b}$ functions yield the full range of colors reproducible by the display, the negative portions of these functions are usually ignored in digital camera designs, which makes the physical reproduction of all colors impossible [8–10]. This mismatch between acquired and displayed image information suggests further study is needed to optimize color matching functions for digital camera color image acquisition. Principal component analysis (PCA) may seem a natural approach to this problem, however, due to orthogonality constraints, PCA basis functions usually contain negative values which complicate physical implementation. Alternatively, the present approach using NMF provides non-negative decomposition which is appropriate and may provide a modeling solution. This seems plausible because the basis functions found by NMF optimized in spectral space provide approximate representations of basis function shapes, wavelength subranges and locations of modes for color matching functions (Fig. 1). Previously published results [1, 7] have provided limited findings towards approximating spectral optimization, however, the present research aims for the accurate approximation of colors, as is needed to fully solve digital capture and color rendering problems.

In this paper we propose a non-negative matrix factorization approach that uses a genetic algorithm (GA) as an optimization technique. Our aim is to find non-negative basis functions that provide accurate color reproduction for a given spectral dataset. Three forms of novel algorithms are developed: non-negative matrix factorization in spectral space; non-negative matrix factorization for color space; and

[†]The author is with Department of Electrical Engineering and Energy Technology, University of Vaasa, P.O. Box 700, FIN-65101 Vaasa, Finland.

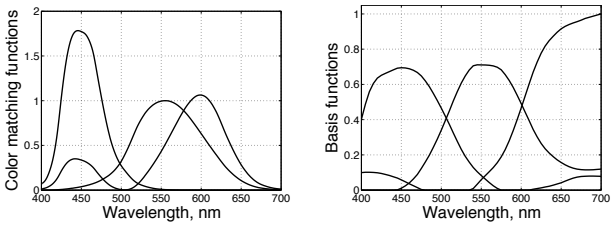
[††]The author is with Institute for Mathematical Behavioral Sciences, University of California, Irvine, CA, USA.

[†††]The author is with Research Center for Frontier Medical Engineering, Chiba University 1-33, Yayoi-cho, Inage-ku, Chiba-shi, Chiba, 263-8522 Japan.

optimization in both spectral and color spaces. Although the separate algorithms can be thought of as integral to a coherent system, for convenience and ease of presentation, we implement and demonstrate each of these forms here as a separate algorithm. Below we test the utility via a series of demonstrations. We first investigate if the algorithm optimized in spectral space achieves similar performance in accuracy as a standard form of NMF. This provides confidence that our novel algorithm performs similarly when compared to a reference algorithm. Next we modify the algorithm optimized in spectral space to achieve new functionality for color space – an extension that a standard NMF is currently not capable of providing. This straightforward modification achieves optimization in color space, greatly increasing the algorithm's accuracy, and demonstrates that the modified variant functions properly. Finally, the algorithm is modified to implement both optimization in color and spectral spaces, demonstrating the success and utility of both implementations.

In the literature genetic algorithms have been utilized for enhancing NMF initialization [11–13]. For example, Clouch and Boukhetala address situations where an algorithm solves a semi-nonnegative matrix factorization for the case where the data matrix $\boldsymbol{X}$ may have both positive and negative signs [11]. These previous results [11] differ from those of the present investigation in which the data matrix is nonnegative. Other results find that evolutionary algorithms for NMF initialization [12] are more computationally demanding than NMF based on nonnegative double singular–value–decomposition (NNDSVD) [14]. By comparison the present study adopts a new SVD-based initialization method that improves upon approaches that previously used NNDSVD [15]. In addition, one investigation previously combined a genetic algorithm and NMF to assess a Japanese Female Facial Expression image dataset, and showed their algorithm has advantages in terms of relative error, sparsity and orthogonality [13]. However, for the kinds of spectral datasets we use here, while some degree of orthogonality is attractive, strong sparsity is undesirable. This is because spectral datasets typically represent curves that are smooth and broadband. Also, earlier investigations [12, 13] that explored genetic algorithms for initialization of NMF algorithms used NMF algorithms which were strictly based on an updating rule. By comparison, the algorithms employed here use closed-form solutions for initialization and the form of NMF employed uses a genetic algorithm. As a result, the new approach presented here can implement optimization in color space, which is an advance when compared to the previous investigations, which were not designed for and, thus, not capable of such an implementation.

The paper is arranged as follows: Methodology is presented in Section 2, the algorithms are described in Section 3, experimental results are given in Section 4, and conclusions are given in Section 5.



a) b)

Fig. 1: Spectral curves. a) CIE 1931 color matching functions of the Standard Observer. b) Normalized basis functions calculated using the standard NMF and optimized in a spectral space.

2. Methodology

Previous work shows the GA is known as a search technique for large spaces and is superior to other optimization methods when requirements for function features like smoothness, continuity and differentiability are violated [16]. In addition, the GA concept is easy to understand and implement, works in noisy environments and supports multiobjective optimization.

The GA contains a population of individuals and assesses the fitness of individuals for each generation. We consider a data matrix $\boldsymbol{X} \in \mathbb{R}^{d\times n}$ and factors $\boldsymbol{W} \in \mathbb{R}^{d\times r}$ and $\boldsymbol{Z} \in \mathbb{R}^{r\times n}$, where d is the number of wavelengths, n is the number of spectral colors, and r is the maximum number of basis functions or the rank of factorization. The number of basis functions can be less or equal to the rank of factorization. We use the GA to define the best subset of variables according to the fitness function:

$$\min_{\boldsymbol{W},\boldsymbol{Z}} \sum_{i=1}^{n} \|\boldsymbol{x}_i - \boldsymbol{W}\boldsymbol{z}_i\|^2 \quad \text{subject to } \boldsymbol{W} \geq \boldsymbol{0};\ \boldsymbol{z}_i \geq \boldsymbol{0}, \tag{1}$$

where $\boldsymbol{x}_i$ and $\boldsymbol{z}_i$ are the columns of the matrices $\boldsymbol{X}$ and $\boldsymbol{Z}$, respectively, and $\boldsymbol{W} \geq \boldsymbol{0}$ and $\boldsymbol{z}_i \geq \boldsymbol{0}$ indicate that all entries of $\boldsymbol{W}$ and $\boldsymbol{z}_i$ are non-negative.

On the other hand, GA provides an effective way to select the best variables for color representation minimizing the color difference. The representation of individuals in this case is illustrated in Fig. 2. The individuals or "chromosomes" are represented by real valued vectors that live in a genotype space. Problem–specific solutions are located in the visual phenotype space as colors. The mapping between the spaces is made using encoding and decoding. The encoding is usually not used while decoding the real valued vectors (spectra), but helps to evaluate individuals according to their fitness function. From a spectral color viewpoint the genotype space corresponds to the physical spectral space, while the phenotype space relates to the physiological color space. Decoding is identical to the spectrum-to-color conversion typically used in spectral imaging [17]. Thus, the real–valued vectors/chromosomes are spectra and the problem–specific solutions achieve color

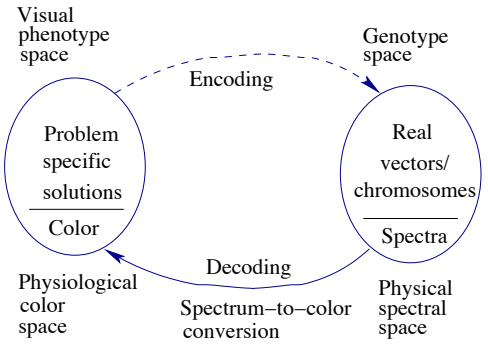


Fig. 2: "Genotype" and "phenotype" spaces and the mapping between the spaces.

difference minimization. Let $\boldsymbol{f}()$ be a decoding function, i.e. a spectrum-to-color conversion function, then $\left\|\Delta E^*_{abi}\right\|^2 = \|\boldsymbol{f}(\boldsymbol{x}_i) - \boldsymbol{f}(\boldsymbol{W}\boldsymbol{z}_i)\|^2$, where ΔE^*_{ab} is a CIELAB color difference: $\Delta E^*_{ab} = \sqrt{(L^*_x - L^*_{Wz})^2 + (a^*_x - a^*_{Wz})^2 + (b^*_x - b^*_{Wz})^2}$. We denote $[L^*_x \;\; a^*_x \;\; b^*_x]^T = \boldsymbol{f}(\boldsymbol{x})$ and $[L^*_{Wz} \;\; a^*_{Wz} \;\; b^*_{Wz}]^T = \boldsymbol{f}(\boldsymbol{W}\boldsymbol{z})$. In this case the fitness function in color space is as follows:

$$\min_{\boldsymbol{W},\boldsymbol{Z}} \sum_{i=1}^{n} \|\boldsymbol{f}(\boldsymbol{x}_i) - \boldsymbol{f}(\boldsymbol{W}\boldsymbol{z}_i)\|^2 \quad \text{subject to } \boldsymbol{W} \geq \boldsymbol{0};\; \boldsymbol{z}_i \geq \boldsymbol{0}. \tag{2}$$

3. Algorithms

Three algorithms based on fitness functions are developed. The first algorithm is for basis functions optimized in a spectral space (GA-NMF-S). The second algorithm is for basis functions optimized in a color space (GA-NMF-C). The third algorithm uses multiobjective optimization for both the spectral and color spaces (GA-NMF-M). The pseudo-code for all these algorithms is as follows:

$t := 0$;
initialize $P(t)$;
evaluate $P(t)$;
while *not terminate* **do**
  $P'(t) :=$ *select-mates*$(P(t))$;
  $P''(t) :=$ *crossover*$(P'(t), pc)$;
  $P'''(t) :=$ *mutate*$(P''(t), pm)$;
  $P''''(t)) :=$ *smooth*$(P'''(t), ps)$;
  evaluate$(P''''(t))$;
  $P(t) := P''''(t)$;
  $t := t + 1$;
end
Algorithm: GA-NMF-i algorithm, i = S,C,M.

Let $P(t)$ denote a population consisting of individuals, and pc, pm and ps the probabilities for operations: "crossover", "mutate", and "smooth", respectively. We use two kinds of populations to define the columns of $\boldsymbol{W}$, i.e. basis functions, and the rows of $\boldsymbol{Z}$. For both populations the sequence of procedures presented in the algorithm is the same and, apart from the operation "mutate", all procedures are identical.

3.1 Initialization

The presented Algorithm is similar to general GA except for the initialization and the operation "smooth". The initialization is important because it can affect algorithm performance. In our earlier study we used a random initialization of $\boldsymbol{W}$ with Gaussian functions [18]. These functions were utilized based on *a priori* knowledge that the resultant basis functions should be bell–shaped and to provide a diversity of population that is needed for evolution. For this purpose Gaussian distributions are more desirable than, for example, piecewise linear functions. However, random initialization of the GA, as well as the standard NMF, leads to variations in observed results. To avoid random initialization influences and to improve performance of the standard NMF a non-negative singular value decomposition (SVD) for initialization has been proposed [15]. We adopted this approach for GA-NMF and modified our operation "initialization", and used singular value decomposition:

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^T,$$

where $\boldsymbol{S}$ is the diagonal matrix of singular values, $\boldsymbol{S} = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$, $\sigma_1 > \sigma_2 > \ldots, \sigma_n$, the columns of matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ are the left and right singular vectors of $\boldsymbol{S}$, respectively, and T is transpose.

Let us assume that the rank r of factorization is less than n, $r < n$, and compute a fraction ($< 90\%$) of the total data variance to be retained. The retained variance in percents is as follows:

$$\frac{\sum_{i=1}^{r} \sigma_i}{\sum_{i=1}^{n} \sigma_i} \times 100\% < 90\%.$$

We then define matrices $\boldsymbol{U}_r = (\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r)$ and $\boldsymbol{S}_r = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$ consisting of the first r vectors and the first r singular values of matrices $\boldsymbol{U}$ and $\boldsymbol{S}$, respectively.

Finally, we initialize matrices

$$\boldsymbol{W}_{init} = |\boldsymbol{U}_r|, \qquad \boldsymbol{Z}_{init} = \left|\boldsymbol{S}_r \boldsymbol{V}^T\right|,$$

where $|\boldsymbol{A}|$ denotes a matrix consisting of the absolute values of the entries of $\boldsymbol{A}$.

The rank r defines the maximum number of basis functions. In experiments we use the number r_b of basis functions, $r_b \leq r$.

The initialization made for the standard NMF can be used to generate only one individual. Therefore, we modify this approach to initialize the whole population. We employ SVD initialization followed by NMF with an update rule called local non-negative matrix factorization (LNMF) [15]:

$$\boldsymbol{Z} \leftarrow \sqrt{\boldsymbol{Z} .* (\boldsymbol{W}^T(\boldsymbol{X}./(\boldsymbol{W}\boldsymbol{Z})))},$$
$$\boldsymbol{W} \leftarrow \boldsymbol{W} .* ((\boldsymbol{X}\boldsymbol{Z}^T)./(\boldsymbol{W}\boldsymbol{Z}\boldsymbol{Z}^T)),$$

where $.*$ and $./$ indicate element-wise multiplication and division, respectively.

We take each even solution of the first 100 iterations of the SVD-NMF algorithm to obtain the population size 50. To determine the population size we follow previous work suggesting that the optimal population size for problems coded as bitstrings is approximately the length of the string [19]. Though we use real valued strings we follow this rule. Hence, the population size, i.e. 50, is approximately defined by the number of wavelengths, i.e. 61.

The solution obtained at iteration 100 does not dominate earlier iterations. The cross-correlation coefficients for the solutions of the MCC dataset at iterations 20 and 100 are as follows: 0.96, 0.98 and 0.97 for short, middle and long wavelengths, respectively. This means that recombination (crossover and mutation) easily achieves a better solution than the best one obtained at initialization.

3.2 Selection, crossover and mutation

The operation "select–mates" makes the tournament selection of parents. In single objective optimization we randomly select the group of 15 individuals from the population. The group size is experimentally defined. Then, within this group the individual with the best fitness value is selected as a parent. For the selection of parents in multiobjective optimization two individuals are used. The selection procedure is given in Section 3.4.

The operation "crossover" is based on uniform crossover with probability $pc = 0.9$.

The operation "mutate" is different for $\boldsymbol{W}$ and $\boldsymbol{Z}$. For individuals of $\boldsymbol{W}$, the gene value w_{ij} is randomly selected and changed using a random value $a \sim \mathcal{N}(0,1)$ as follows: $w_{ij} = w_{ij} + 0.0005a$. Similarly for z_{ij}, $z_{ij} = z_{ij} + 0.2a$. When the value w_{ij} or z_{ij} become less than zero it is set to zero. The selected value for probability of mutation is $pm = 1/61$.

The designed algorithms have a deterministic initialization, but selection is still random. Therefore the results have small variations at each algorithm run which can be neglected. For example for three datasets used in this study and spectral optimization described in Section 4, the confidence interval of the average MSE is $[0.1386 \times 10^{-2}, 0.1414 \times 10^{-2}]$ (MCC dataset), $[0.8377 \times 10^{-2}, 0.8623 \times 10^{-2}]$ (paint dataset) and $[0.1829, 0.1869]$ (Munsell dataset) at confidence level equal to 95%.

3.3 Smoothing

Crossover and mutation can produce abrupt changes in the values of a chromosome's neighboring genes, leading to multi–peaked basis functions. To solve this problem we introduce a smoothing operation "smooth" for columns of $\boldsymbol{W}$ and for rows of $\boldsymbol{Z}$ with a simple averaging window comprising three elements. Testing shows that a good probability of smoothing is about 10 times less than the probability of crossover, i.e. $ps = 0.1$.

3.4 Evaluation and fitness functions

For algorithms GA-NMF-S and GA-NMF-C, fitness functions are identical to the objective functions Eqs. (1) and (2), respectively.

In multiobjective optimization (GA-NMF-M) a Pareto ranking approach is utilized [20]. The objective value of each solution is not assigned in this case. The population is ranked according to a dominance rule in the objective space. The operation "evaluate" is replaced by three operations: "evaluateS", "evaluateC" and "rank":

$E_s(t) := evaluateS\ (P(t))$;
$E_c(t) := evaluateC\ (P(t))$;
$rank(E_s(t), E_c(t))$;

The operations "evaluateS" and "evaluateC" calculate the objective functions $E_s(t)$ (Eq. (1)) and $E_c(t)$ (Eq. (2)), respectively. The solutions are evaluated using the operation "rank" which calculates the fitness value based on the solution rank in population [20, 21].

First, the Pareto ranking technique calculates a set of nondominant fronts F_i, where $i = 1, 2, \ldots$, for the whole population, where lower ranks correspond to better solutions. Hence, F_1 is a Pareto front. Second, we use the ranking technique for solutions belonging to the particular front. We adopt the crowding distance approach [20] and modify it introducing a more efficient hypervolume measure [22]. For three sequential solutions/points located in the front, the hypervolume measure for the central point is $v = d_c d_s$, where d_c and d_s are distances between the two extreme points related to E_c and E_s, respectively.

For the parent selection we randomly take two solutions. If solutions are from different nondominant fronts, the solution with the lowest rank is selected. Otherwise, the solution with the higher hypervolume measure wins.

3.5 Elitism

Here, "Elitism" relates to the problem of keeping good solutions during optimization. In single objective optimization we visually monitor the objective function value over time and select algorithm parameters achieving a good convergence rate. Therefore, we do not maintain elitist solutions in the single objective optimization case.

For multiobjective optimization we use the elitism strategy combining the old population and the offspring. Then, we apply deterministic selection keeping the best solutions to obtain the new population with the size of 50.

4. Experiments

The experiments were conducted with stimulus measures using optimization in two spaces: spectral and color spaces. The measured number of wavelength of spectra was taken at 5nm intervals in the humanly visible range (400-700nm). The number of basis functions tested was three.

Four light sources (D65, Halogen, LED and Xenon) were utilized in tests. Three datasets of measured reflectance spectra were used in experiments (Tab. 1). For the paint dataset we increased the number of wavelengths originally taken at 10nm from 31 to 61 using linear interpolation.

NMF can produce ambiguous solutions with respect to factors $\boldsymbol{W}$ and $\boldsymbol{Z}$. To avoid this problem, two normalizations of calculated basis functions were completed. That is, we first divided all elements in each column of $\boldsymbol{W}$ by their sum [5], and subsequently normalized all basis functions to make their total maximum equal unity.

Table 1: Spectral datasets. The first dimension of "Size" equals number of wavelengths (#w) and the second one equals the number of samples (#s).

Dataset	Size=#w×#s
MCC	61×24
Munsell color dataset [23]	61×1269
Paint dataset [24]	61×91

4.1 Spectral space

Non-negative matrix factorization was implemented in the spectral space. The Algorithm used fitness function according to Eq. (1) and the initialization described in Section 3.

The parameters for GA were as follows. The individual length for columns of $\boldsymbol{W}$ was 61 (number of wavelengths #w). The individual length for rows of $\boldsymbol{Z}$ corresponded to the number of samples (#s) (Tab.1). The number of iterations was 200,000, although for the Munsell set it was 1,000,000. The population size was 50 for all tests.

We initially compare GA-NMF-S with a standard NMF algorithm which can be divided into several classes: Multiplicative update algorithms, gradient descent algorithms, and alternating least squares (ALS) algorithms. For these algorithms the statements about convergence (global or local) have not been proven, although it has been suggested that the ALS algorithm will converge to a local minimum in certain special cases [3]. The multiplicative update algorithms are the first well–known NMF algorithms which often converge in practice. They are, generally speaking, known to be a standard baseline against which any new algorithm is compared. Therefore, a multiplicative update algorithm was selected for comparison with GA. Note that convergence of GA has not been proven, and it is known that GA can be found to converge to a local minimum. However, we have confidence in the present solutions due to their wellformedness compared to known chromatic response functions of human observers, and because GA algorithms in general are widely used to obtain high-quality solutions for difficult and complex optimization problems.

For comparison we give the result for MCC using the standard NMF algorithm, 200,000 iterations (Fig. 1b) [4]. The result for GA-NMF-S is shown in Fig. 3a. The results are similar. In addition, the MSE values for each spectral color are given (Tab. 2). The GA-NMF results are better than those of the standard NMF.

Table 2: Spectral approximation. MSE for standard NMF, GA-NMF-S, and GA-NMF-M. The table cell order corresponds to the order of color patches of MCC.

Standard NMF, average: 0.0031					
0.0005	0.0035	0.0013	0.0010	0.0037	0.0049
0.0045	0.0011	0.0031	0.0017	0.0032	0.0052
0.0014	0.0018	0.0045	0.0049	0.0040	0.0035
0.0125	0.0055	0.0022	0.0008	0.0002	0.0000
GA-NMF-S, average: 0.0014					
0.0004	0.0006	0.0004	0.0007	0.0028	0.0037
0.0035	0.0004	0.0005	0.0021	0.0008	0.0033
0.0009	0.0013	0.0024	0.0003	0.0031	0.0034
0.0012	0.0004	0.0003	0.0001	0.0000	0.0000
GA-NMF-M, average: 0.0017					
0.0004	0.0006	0.0005	0.0011	0.0034	0.0042
0.0048	0.0005	0.0007	0.0025	0.0006	0.0051
0.0012	0.0014	0.0030	0.0007	0.0040	0.0046
0.0013	0.0005	0.0004	0.0003	0.0001	0.0001

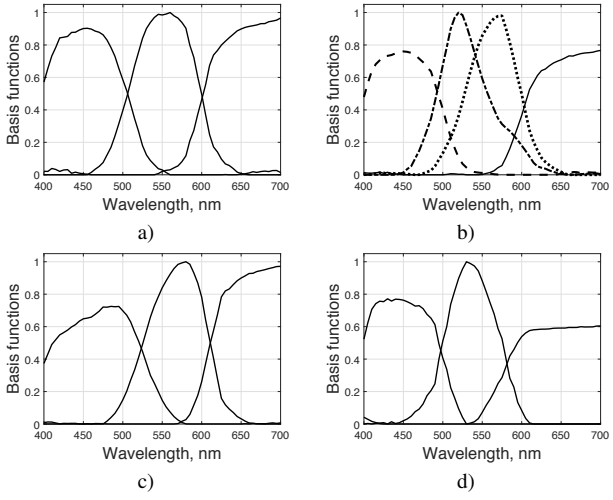


Fig. 3: Algorithm GA-NMF-S. a) Three normalized basis functions for MCC. b) Four normalized basis functions for MCC. c) Three normalized basis functions for the paint dataset. d) The three normalized basis functions for the Munsell dataset.

Algorithms were implemented as Matlab routines installed on a 2.5GHz laptop computer running Windows operating system. The computational time for standard NMF obtained 101s versus 220s for the GA-NMF-S. Below we modify the GA-NMF algorithm to obtain a new functionality to work in color space which the standard NMF does not provide. The standard NMF is based on an update rule working in the physical spectral space and is not capable of working in color space. It will be shown that the optimal basis functions derived in color space approximate color appearance similarity significantly better.

Though our main purpose is to study three-band (primary or filter) systems we additionally test the potential of our method for a four-band system. We note that SVD initialization defines the rank of matrix factorization [15]. According to this approach the rank cannot be higher than four,

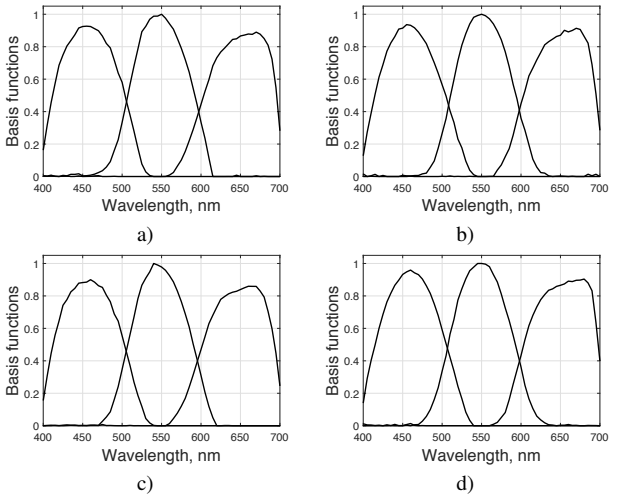


Fig. 4: Normalized basis functions calculated using GA-NMF-C for MCC and different light sources: a) D65, b) Halogen, c) LED, and d) Xenon.

i.e. four basis functions, for MCC. The result is shown in Fig. 3b. In addition, we tested other datasets. The results are shown in Figs. 3c and d. The results are somewhat similar to MCC colors. The different results for the paint and Munsell dataset can be explained by different color measurements presented in these sets.

4.2 Color space

The GA-NMF gives a simple and straightforward way for matrix factorization optimized for color difference, which is not an option with standard NMF.

We utilize GA-NMF-C with a fitness function based on color difference (Eq. (2)) and initialization described in Section 3. The rest of the GA parameters are identical to that of GA-NMF-S. The spectrum-to-color conversion is used for decoding [17]. For conversion we use the illuminants: D65, Halogen, LED and Xenon and CIE 1931 color matching functions. The results are obtained using 200,000 iterations. The results are presented in Fig. 4 and Tab. 3.

From the basis functions and color matching functions in Figs. 1a, 3a, and 4a, and Tab. 3, one can clearly see that the basis functions optimized in color space are much more accurate than in spectral space. For MCC colors ΔE_{avg} is reduced about six times (Tab. 3). The basis functions are quite bell shaped, and better define the wavelength subranges and modes of the curves. For the long-wavelength curve the result is still not good but this curve at least approximately defines the subrange of wavelengths. Surprisingly, the results does not change much if light is varied (Fig. 4a-d).

In addition, we tested the paint dataset and the Munsell dataset using the GA-NMF-C algorithm given 200,000 iterations and 1,000,000 iterations, respectively. Assuming the same light source (D65 illuminant) for both datasets, we observed the basis functions tails for the Munsell dataset decayed more rapidly (Fig. 5a) compared to those of curves in Fig. 3d. Moreover, the curves are less smooth. Similar result

Table 3: Color approximation, light source D65. The color difference CIELAB ΔE. The table cell order corresponds to the order of color patches of the MCC.

GA-NMF-S, average: 6.47					
2.10	1.55	2.20	7.90	6.92	12.85
14.24	3.62	4.18	8.75	7.62	11.58
10.93	13.45	11.86	3.25	5.09	16.28
2.22	2.49	2.20	1.96	1.36	0.77
GA-NMF-C, average: 1.14					
2.62	0.10	0.14	1.46	1.32	0.64
0.67	1.15	1.72	0.80	1.14	1.42
0.96	0.54	0.80	0.52	1.21	2.85
0.43	1.44	1.31	0.68	0.96	2.37
GA-NMF-M, average: 2.99					
2.07	1.51	0.90	2.82	3.97	6.09
4.98	0.84	2.00	4.23	2.33	5.01
1.32	4.68	4.38	2.71	3.12	6.00
2.37	2.00	1.39	1.49	1.59	4.05

was obtained for the paint dataset. For GA-NMF-S (paint dataset) $\Delta E_{avg} = 15.05$ and for GA-NMF-C $\Delta E_{avg} = 7.56$. For GA-NMF-S (Munsell dataset) $\Delta E_{avg} = 12.53$ and for GA-NMF-C $\Delta E_{avg} = 8.52$.

4.3 Multiobjective optimization

We also investigated the advantage of GA in multiobjective optimization (GA-NMF-M). In this experiment we search for optimal basis functions for both spectra and colors using criterion based on the Pareto front described in Section 3. The results are presented in Fig. 5b, Tabs. 2, and 3. The results represent a trade-off between two marginal cases: spectral and color optima. The algorithm is less accurate than corresponding single objective algorithms in color space.

We suppose that the solution given by the GA-NMF-S algorithm is close to optimal. The approximation obtained from the multiobjective algorithm can be considered good because the ratio indicating fit, defined as the average MSE of the GA-NMF-S algorithm over the average MSE of multiobjective optimization, is 0.82, i.e. close to unity. In addition, the average MSE of multiobjective optimization is better than the average MSE of the standard NMF (Tab. 2). The best rank solution of the GA-NMF-M algorithm gives $E_c = 71.85$ and $E_s = 0.042$ at iteration 200,000 while $E_c = 1.34 \times 10^4$ and $E_s = 17.7$ at iteration 300 which are far from the optimum. Based on these findings we consider the multiobjective optimization to be efficient.

The above results give us confidence that the multiobjective algorithm can be effectively used in digital image archiving for the purposes of employing one set of basis functions, simplifying the task of deriving two sets representing spectral and color space solutions.

5. Conclusions

We developed new algorithms for NMF using a GA approach. One algorithm presented, GA-NMF-S, finds optimal basis functions for a spectral dataset in the spectral do-

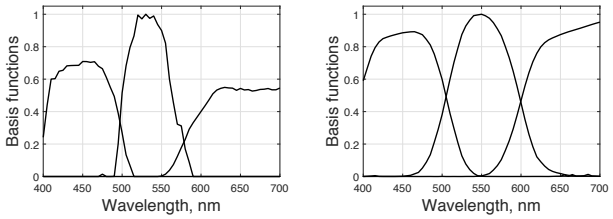


Fig. 5: Normalized basis functions. a) Algorithm GA-NMF-C for Munsell dataset. b) Algorithm GA-NMF-M for MCC.

main. That algorithm's performance was shown to be similar to the standard NMF algorithm. A second algorithm was introduced, GA-NMF-C, that was shown to provide optimal basis functions for a color space. To the best of our knowledge this is a new result. The basis functions from GA-NMF-C were shown to be bell shaped and more accurately capture the mode locations of the spectral curves. In addition, the second algorithm accurately approximates colors. Finally, multiobjective optimization in both spectral and color spaces was implemented using the third algorithm modification GA-NMF-M. Investigations of the latter implementation found that the multiobjective algorithm can be used in digital image archiving for the purposes of employing one set of basis functions, and thereby simplifying the task of deriving two sets representing spectral and color space solutions.

An important feature of the present approach is that the procedures are both general and scalable, in the sense that they are not bound to the use of CIE 1931 color space, per se, or to the choice of particular R,G,B display primaries, or constrained by the 5nm measurement intervals present in the wavelength measurements in our spectral datasets. Thus, the methods described here can be generalized and make use of any color appearance spaces that has some form of color difference metric, and across a reasonable set of naturalistic illuminants. Moreover, as seen by the demonstration of 4 basis functions approximated using these techniques, the present approach can be used to evaluate and approximate color processing dimensions that differ from that commonly used when developing three-parameter device profiles as a model of the color perception of a standard normal human observer.

Acknowledgments

Jameson was in part supported by an award from the National Science Foundation (SMA-1416907, K.A. Jameson, PI).

References

[1] A. Kaarna, K. Tamura, S. Nakauchi, and J. Parkkinen, "Non-negative bases for spectral color sets," Proc. 2nd Int. Workshop on Image Media Quality and its Applications, Chiba, Japan, pp. 333–343, 2007.

[2] Y. Miyake and V. A. Bochko, "Multispectral imaging," in Handbook of Digital Imaging, Ed. by M. Kriss, Wiley, 2015.

[3] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," Comp. Stat. Data Anal., vol. 52, no. 1, pp. 155–173, 2007.

[4] P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," Journal of Machine Learning Research, vol. 5, pp. 1457–1469, 2004.

[5] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization, " Nature, vol. 401, no. 6755, pp. 788–791, 1999.

[6] P. Paatero and U. Tapper, "Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values," Environmetrics, Vol. 5, 111–126, 1994.

[7] A. Andriyashin, "Non-negative bases in spectral image archiving," Ph.D. Thesis, University of Eastern Finland, Joensuu, Finland, 2011.

[8] B. Hill, "High quality image reproduction: The multispectral solution," Proc. 9th Int. Symposium on Multispectral Colour Science and Application, Taipei, Taiwan, pp. 1–7 ,2007.

[9] M. Kriss, "Color reproduction in digital cameras," in Handbook of Digital Imaging, Ed. by M. Kriss, Wiley, 2015.

[10] W. K. Pratt, Digital Image Processing, John Wiley & Sons, Inc., New York, NY, 1978.

[11] M. Chouh and K. Boukhetala, "Semi-nonnegative matrix factorization algorithm based on genetic algorithm initialization," Int. J. Machine Learning and Computing, Vol. 6, No. 4, pp. 231-234, 2016.

[12] A. Janecek and Y. Tan, "Using population based algorithms for initializing nonnegative matrix factorization," in Advances in Swarm Intelligence, Lecture Notes in Computer Science, Vol. 6729, pp. 307-316, 2011.

[13] M. Rezaei and R. Boostani, "Using the genetic algorithm to enhance nonnegative matrix factorization initialization," Expert Systems, Vol. 31, No. 3, pp. 213-219, 2014.

[14] C. Boutsidis and E. Gallopoulos, "SVD based initialization: A head start for nonnegative matrix factorization," Pattern Recogn. Vol. 41, No. 4, pp. 1350-1362, 2008.

[15] H. Qiao, "New SVD based initialization strategy for non-negative matrix factorization," Pattern Recogn. Letters, vol. 63, pp. 71–77, 2015.

[16] M. Mitchell, An Introduction to Genetic Algorithms, Cambridge, MA: MIT Press, 1996.

[17] Colorlab Toolbox for Matlab by the University of Joensuu Color Group (http://spectral.joensuu.fi)), Accessed: May 1, 2016.

[18] V. A. Bochko, Y. Miyake, and J. T. Alander, "Non-negative matrix factorization for Macbeth ColorChecker spectral colors using genetic algorithm," Proc. 8th Int. Workshop on Image Media Quality and its Applications, Nagoya, Japan, pp. 13–17, 2016.

[19] J. T. Alander, "On optimal population size of genetic algorithms," In Computer Systems and Software Engineering, 6th Annual European Computer Conference, P. Dewilde and J. Vandewalle, Eds., pp. 65–70, 1992.

[20] A. Konak, D. W. Coit, and A. E. Smith, "Multi-objective optimization using genetic algorithms: A tutorial," Reliability Engineering & System Safety, vol. 91, pp. 992–1007, 2006.

[21] D.E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Reading, MA: Addison-Wesley, 1989.

[22] M. Emmerich, N. Beume, and B. Naujoks, "An EMO algorithm using the hypervolume measure as selection criterion," In Evolutionary Multi-Criterion Optimization, C. A. Coello Coello, A. H. Aguirre, and E. Zitzler, Eds., Lecture Notes in Computer Science, Vol. 3410, pp. 62-76, 2005.

[23] Spectral Database, University of Joensuu Color Group, (http://spectral.joensuu.fi/), Accessed: May 1, 2016.

[24] J. Tajima, M. Tsukada, Y. Miyake, et al., "Development and standardization of a spectral characteristics data base for evaluating color reproduction in image input devices," Proceedings of SPIE, vol.

3409, pp. 42–50, 1998.

Vladimir A. Bochko received a Dr. Sc. in Engineering degree from St. Petersburg State Electrotechnical University in 1987. He worked in several research positions in St. Petersburg State Electrotechnical University, Imperial College, and University of Joensuu. He worked in Lappeenranta University of Technology as an Assistant Professor. He was the visiting researcher of the Department of Information and Image Sciences, Chiba University, Japan in 2005 (4 months) and 2006 (2 months). He is now with the University of Vaasa as an Assistant Professor.

Kimberly A. Jameson received her PhD in Psychology in 1989 from Cognitive Sciences at the University of California, Irvine. During 1995-2003 she was Assistant Professor of Psychology at UC San Diego. In 2003 she joined the Institute for Mathematical Behavioral Sciences at UC Irvine, where she is currently Project Scientist. Her empirical and theoretical work includes research on color perception and photopigment opsin genetics; mathematical modeling of color category evolution among communicating artificial agents; and individual variation and universals in human color cognition and perception. She directs the Color Cognition Laboratory at UC Irvine, and is the lead researcher responsible for development and public-access dissemination of the ColCat cross-cultural color categorization digital archive.

Toshiya Nakaguchi received B.E., M.E., and Ph.D. degrees from Sophia University, Tokyo, Japan in 1998, 2000, and 2003, respectively. He was a Research Fellow supported by Japan Society for the Promotion of Science from 2001 to 2003. He moved to the Department of Information and Image Sciences, Chiba University in 2003, as an Assistant Professor. From 2006 to 2007, he was a Research Fellow in Center of Excellence in Visceral Biomechanics and Pain, in Aalborg Denmark, supported by CIRIUS, Danish Ministry of Education. He moved to the Department of Medical System Engineering, Chiba University in 2010 as an Associate Professor. He moved to Center for Frontier Medical Engineering, Chiba University in 2013, as an Associate Professor. He is currently a Professor in Center for Frontier Medical Engineering, Chiba University since 2016. His research interests include medical engineering, color image processing, computer vision and computer graphics.

Yoichi Miyake received Ph.D. from Tokyo Institute of Technology in 1978. During 1989-2009 he was the Professor of Chiba University. From 2009 to 2011 he was the Research Professor of Chiba University. He was the postdoctoral fellow of Swiss Federal Institute of Technology (ETHZ) in 1978 and 1979. He was the guest professor of University of Rochester in 1997. He received Charles E. Ives Award (paper award) from IS&T in 1991, 2001 and 2005. He was named as Electronic Imaging Honoree of the year in 2000 from SPIE and IS&T. He received Edwin H. Land Medal from OSA and IS&T in 2012 as an inventor of spectral imaging endoscopes. He received CAT (Color Association of Taiwan) award in 2014. Beginning with 2009 he became the Emeritus Professor. He is currently with Tokyo Polytechnic University as a managing director. He published many books and original papers on image processing. He is known as a pioneer of spectral image processing.

Jarmo T. Alander received the B.Sc., M.Sc., and Lic. Phil. degrees in Physics from University of Helsinki in 1978, 1983, and 1999, respectively and M.Sc., Lic. Tech., and Dr. Sc. in Engineering from Helsinki University of Technology in 1985, 1989, and 2001, respectively. He has worked in several assistant positions in both University of Helsinki and Helsinki University of Technology from 1978 until 1993. From 1993 he has held the professorship in production automation in University of Vaasa. He has published many original papers in genetic algorithms and various other automation topics related to signal and image processing.