The Role of Color and Contrast in Facial Age Estimation

Hamdi Dibeklioğlu[1,2], Theo Gevers[1], Marcel Lucassen[1], and Albert Ali Salah[3]

[1]Intelligent Systems Lab Amsterdam, University of Amsterdam, The Netherlands
{h.dibeklioglu,th.gevers,m.p.Lucassen}@uva.nl
[2]Pattern Recognition & Bioinformatics Group, Delft University of Technology, The Netherlands
[3]Department of Computer Engineering, Boğaziçi University, Istanbul, Turkey
salah@boun.edu.tr

Abstract. Computer based methods for facial age estimation can be improved by incorporating experimental findings from human psychophysics. Moreover, the latter can be used in creating systems that are not necessarily more accurate in age estimation, but strongly resemble human age estimations. In this paper we investigate the perceptual hypothesis that contrast is a useful cue for estimating age from facial appearance. Using an extensive evaluation paradigm, we establish that using a perceptual color space improves computer's age estimation, and more importantly, using contrast-enabled features results in estimations that are more correlated to human estimations.

Keywords: Age estimation, age perception, facial contrast, facial color

1 Introduction

Age estimation is a perceptual task we perform automatically and often unconsciously, as a regulator of social interactions. Age estimation for youngsters is important to estimate cognitive capacities, whereas in general the age information would provide historical information usable in social contexts (e.g. "You certainly would not remember the time Commoder 64 was popular."). In many cultures, older individuals are accorded a certain respect associated with the age, and simultaneously, direct inquiry about a person's age is often considered inappropriate. The inevitable result is that the age is estimated from available cues, such as the appearance of the face, the tautness of skin and the existence of wrinkles, the color of the hair, the tone of voice, the manner of speaking, perhaps even the choice of clothing. It can be said that the human perceptual system is quite adept at making guesses about a person's age.

In this paper, we investigate the perceptual hypothesis that contrast is a useful cue for estimating age from facial appearance. Computer estimation of age from facial appearance has several applications, and it is important to establish reliable cues for this problem. While there is some evidence in perception studies that humans use contrast cues successfully for this task [14], it is known that many factors affect human perception of age: People are better at estimating age of younger faces or individuals that look like themselves, they are affected by the gender, attractiveness and expression of the estimated face, as well as biased by hair color, contextual cues, and such [20]. Subsequently, an experimental approach is necessary to verify this hypothesis. In this paper, we describe a set of contrast features, and use a state of the art age estimation pipeline to test their usefulness for this problem. We report our results on the publicly available UvA-NEMO database with 400 subjects of ages 8-76 [2]. We establish that 1) using a perceptual color space improves computer's age estimation, 2) enabling contrast features marginally improves the results, although the improvement is more marked for approaches that process grayscale images, and more importantly, 3) using contrast-enabled features results in age estimations that are more correlated to human estimations.

Developing age estimation systems that model human estimation (rather than trying to estimate the true age) has not received much attention in the literature, but such systems are important for certain applications. One example is cosmetics, where the perceived age can be significantly reduced, hence comparisons are more meaningful with perceived age, rather than true age. Another example is the assessment of child exploitation crimes (e.g. child pornography), where an investigator gives a decision about the age of the child by inspecting visual images, and sometimes stakes his or her reputation on a decision, which is difficult to make [10]. In this case, a computer system that approximates the human age estimation can provide objective justification for such decisions.

2 Related Work

2.1 Psychophysical Studies

The few psychophysical studies on age estimation from face images suggest that contrast information from specific face regions and color distribution are indicative on the estimated age. Most studies employ digital manipulation of face images (predominantly females). In [4] this leads to the finding that removal of skin surface topography cues (such as fine lines and wrinkles), but preservation of skin color information, resulted in a decrease of estimated age of about 10 years compared with the age judgments of unmodified faces. In contrast, digital smoothing of facial discoloration resulted in a decrease of perceived age of 1 to 5 years.

In [1] the perceived age of male faces was studied using digital manipulation of shape and color information. While the authors could change the perceived age with color manipulation of individual pixels, they reasoned that this effect was not due to enhanced contrast or color saturation. In [8], skin wrinkling, hair graying and lip height were significantly and independently associated with how old a woman looks for her age. In a study on faces of Caucasian women [11], it was shown that the most important attributes to estimate age are eyes, lips and skin color uniformity. Another study on female faces [14] performed on the CIELAB color space indicates that faces with greater a* (red-green) contrast around the mouth, greater luminance contrast around the eyes, or greater luminance contrast around the eyebrows were judged to be significantly younger. These studies also point out to the importance of color information for age estimation.

2.2 Computer estimation of facial age

In contrast to the findings reported in the previous section, the majority of computer based facial age estimation methods assume gray-scale images. This is partly the case because of the nature of the major benchmarking databases (such as FG-NET [19] and MORPH [16]), which collect old and new photographs of individuals, and consequently have varying degrees of color information in them. In this work, we partly mitigate this by exploring color information in our experiments on the UvA-NEMO database [2], which uses a controlled lighting setup. This database has a large number of subjects and a wide age range (8-76), but it does not allow longitudinal inspections of individuals.

The most important cues used in age classification are appearance-based, most notably the cranio-facial development, which instigated a host of methods that simulate the evolution of facial aging for analysis and synthesis [7, 18], and wrinkles formed on the face due to deformations in the skin tissue [21, 22].

The first class of methods apply subspace projection and manifold embedding techniques to find trajectories of age progression for a given individual. In [6] probabilistic kernel principal component analysis is used for this purpose. The second class of methods apply robust feature extraction approaches that are known to work well in face analysis, and treat the problem as a classification or regression task. In [21] Gabor wavelet features and local binary patterns were used successfully. Good surveys of the facial age estimation field can be found in [15] and [5].

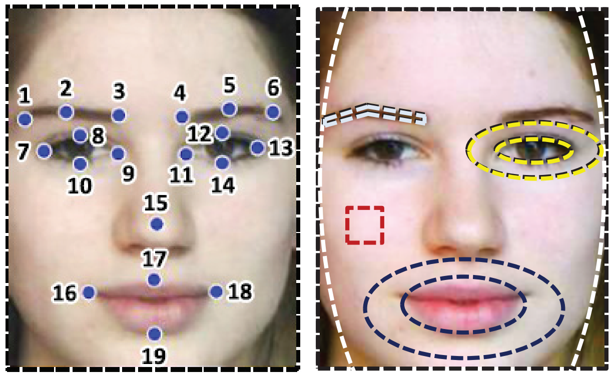
3 Method

In this study, we propose the use of facial contrast with appearance information for automatic age estimation. The input images are assumed to have a moderately frontal face. The flow of the system can be summarized as follows. Initially, 19 facial landmarks are automatically located in the images. Then, by using the detected landmarks, size and rotation of faces are normalized, the regions of interest are cropped, and facial contrast features are extracted. To describe the facial appearance, uniform Local Binary Patterns (LBP) are computed on the input images. Finally, appearance and contrast features are fused to train/test Support Vector Machine (SVM) regressors.

3.1 Features

In the proposed system, facial appearance and contrast features are extracted from images and fused to improve age estimation accuracy as well as increasing the correlation between human perception and automatic estimation of ages. We use CIELAB color space in addition to gray-scale and RGB space, since it was designed as a perceptually uniform color space. It consists of an achromatic lightness channel (L*) and two color opponent channels a* (red-green) and b* (yellow-blue). In approximation, equal distances between two points in this space are also perceptually equal.

Before feature extraction, faces are normalized (with respect to scale and rotation) and regions of interest for facial contrast analysis are cropped using 19 facial landmarks (eyebrow corners/centers, eye corners, center of upper/lower eyelids, nose tip,



(a) (b)

Fig. 1. (a) Used facial landmarks with their indices and (b) the regions of interest on an aligned/cropped face

lip corners, center of upper/lower lips, see Fig. 1(a)). These landmarks are automatically detected using the method proposed in [3]. This method models Gabor wavelet features of a neighborhood of the landmarks using incremental mixtures of factor analyzers and enables a shape prior to ensure the integrity of the landmark constellation. It follows a coarse-to-fine strategy; landmarks are initially detected on a coarse level and then fine-tuned for higher resolution.

Facial Contrast Features. To analyze and describe the facial contrast, we extract a set of features from eyebrows, eyes, lips, and whole face. First of all, eye centers are computed as middle points between inner and outer eye corners as $c_1 = \frac{l_7+l_9}{2}$ and $c_2 = \frac{l_{11}+_{13}}{2}$, where l_i shows 2D coordinates of landmarks. Then, the roll rotation of the face is estimated as $R_{\text{roll}} = \arctan\left(\frac{c_{y,2}-c_{y,1}}{c_{x,2}-c_{x,1}}\right)$, where $c_{x,i}$ and $c_{y,i}$ denote x and y values of center points c_i, respectively. Using the estimated rotation the pose is normalized to frontal face.

After normalization of rotation, face is cropped as shown in Fig. 2. Then, interocular distance d_{io} (Euclidian distance between eye centers) is calculated and the face is scaled with a factor of $80/d_{\text{io}}$. Resultant normalized face image has a resolution of 200×160 pixels.

When the face is normalized, regions of interest are automatically determined, as shown in Fig. 1(b), using landmarks. Regional patches are adapted and modified from the age perception study of Porcheron et al. [14], where eyes, eyebrows, lips, and surrounding areas of those are manually annotated. In this study we automatically crop inner and surrounding regions of eye/lip by fitting an ellipse on the related landmarks. Inner eyebrow regions are cropped using the dilated lines on the eyebrow landmarks. Patches cropped on the cheeks are used as surrounding skin for eyebrows to cope with varying thickness of eyebrows and possible hair occlusions on the forehead. Surrounding regions define the skin, where the inner regions define feature areas. The boundary

Table 1. Definitions of the extracted features

Feature	Definition
Regional Mean Contrast:	$\frac{\text{mean}(I_{\text{skin}})-\text{mean}(I_{\text{in}})}{\text{mean}(I_{\text{skin}})+\text{mean}(I_{\text{in}})}$
Regional Median Contrast:	$\frac{\text{median}(I_{\text{skin}})-\text{median}(I_{\text{in}})}{\text{median}(I_{\text{skin}})+\text{median}(I_{\text{in}})}$
Inner (Michelson) Contrast:	$\frac{\max(I_{\text{in}})-\min(I_{\text{in}})}{\max(I_{\text{in}})+\min(I_{\text{in}})}$
Surrounding (Michelson) Contrast:	$\frac{\max(I_{\text{skin}})-\min(I_{\text{skin}})}{\max(I_{\text{skin}})+\min(I_{\text{skin}})}$
Inner Smoothness:	$\text{std}(I_{\text{in}})$
Surrounding Smoothness:	$\text{std}(I_{\text{skin}})$
Smoothness Rate:	$\frac{\text{std}(I_{\text{in}})}{\text{std}(I_{\text{skin}})}$

of face is also defined to compute global face contrast. Feature area for the face is the combination of the inner eyebrow, eye, and lip regions. Area lays between the face boundary and inner face regions forms the skin area of the whole face.

Let I_{in} and I_{skin} denote the one-channel (such as gray, R, G, B, L*, a*, b*) color values of inner and surrounding area of the related facial region, then facial contrast features can be defined as given in Table 1, where std denotes the standard deviation. The regional contrasts are defined as an adapted version of the Michelson contrast [9]. Note that the defined features are extracted separately for each of the face, eye, eyebrow, and lip regions. For eyes and eyebrows, average of the left and right side regions are used. Additionally, Michelson contrast and smoothness (std) features are extracted for the area within the face boundary. As a result, a 30 dimensional contrast feature vector is composed for the related color channel.

Appearance Features. To describe the facial appearance, uniform *Local Binary Patterns* (LBP) are used. The original LBP operator, which is proposed by Ojala *et al.* [12], takes the intensity value of the center pixel as threshold to convert the neighborhood pixels to a binary code. Computed binary codes describe the ordered pattern of the center pixel. This procedure is repeated for each pixel on the image and the histogram of the resultant 256 labels can then be used as a texture descriptor. In [13], Ojala *et al.* show that the vast majority of the *Local Binary Patterns* in a local neighborhood contain at most two bitwise transitions from 0 to 1 or 1 to 0, which is called a uniform pattern. Therefore, during the computation of the histograms, the size of the feature vector can be significantly reduced by assigning different bins for each of the 58 uniform patterns and one bin for the rest.

Each face is divided into 7×5 non-overlapping (equally-sized) blocks and uniform LBP descriptors are computed on each block (see Fig. 2). 8 neighborhood pixels (on a

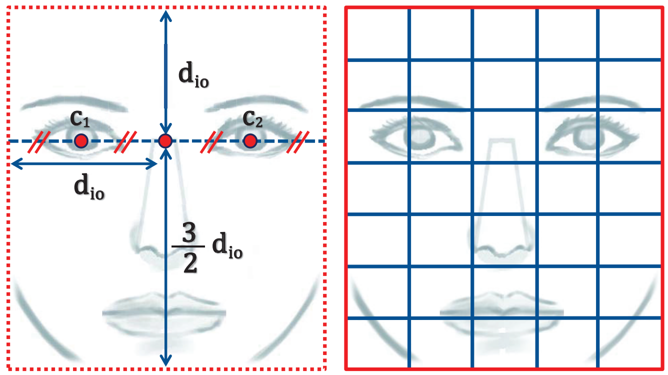


Fig. 2. (a) Scaling/cropping of a face image, and (b) the defined 7×5 blocks to extract appearance features

circle with a radius of 1 pixel) are used to extract the uniform LBP features. All these features are concatenated to form the appearance feature vector. Resultant appearance feature vector is $7 \times 5 \times 59 = 2065$ dimensional per color channel.

3.2 Classification

Extracted contrast and appearance features are concatenated for each color channel. Finally, features for the related color channels are fused. As a result, we have 2095 dimensional feature vectors for gray-scale, where the number of combined features for RGB and CIELAB is $2095 \times 3 = 6285$. Then these features are fed to SVM regressors for age estimation. In order to optimize the SVM configuration, different kernels (linear, polynomial, and radial basis function (RBF)) with different parameters (size of RBF kernel, degree of polynomial kernel) are tested on the validation set and the configuration with the minimum validation error is selected. The test partition of the dataset is not used for parameter optimization. The resulting estimation of the age is given as an integer with 1 year resolution.

4 Experimental Data

4.1 Face Images

To evaluate our system and assess the reliability of facial contrast and appearance information for age estimation, we extract and use the initial neutral frame of each video in freely available UvA-NEMO Smile Database [2] (see Fig. 3). The database has 1240 smile videos from 400 subjects (185 female, 215 male). Ages of subjects vary from 8 to 76 years. Videos were recorded (in RGB) with a resolution of 1920×1080 pixels at a rate of 50 frames per second under artificial D65 daylight illumination. Additionally, a color chart is present on the background of the videos for color normalization. Number of videos per subject varies from 1 to 4.

Fig. 3. Sample face images from the UvA-NEMO Smile Database

4.2 Age Perception

We gathered perceived ages for a subset of the neutral face images extracted from the UvA-NEMO Smile Database. The original recordings show a MacBeth Color Rendition chart. Using the black and white patches of the lightness scale in the MacBeth chart, we color corrected the images to have the same lightness values (L* in CIELAB color space), i.e. L*=96 for the white patch and L*=21 for the black patch, in the sRGB color profile [17]. In total, 84 face images were used from 42 male and 42 female, mainly Caucasian. Actual ages range from 8 to 76 years with a distribution similar to that of whole database. The face images were presented on a Eizo ColorEdge CG211 monitor which was calibrated to the sRGB standard.

Twenty-four participants, 14 male and 10 female, rated the perceived age of the (unknown) faces. All subjects had normal color vision as confirmed by the HRR color vision test and had normal or corrected-to-normal visual acuity. Participants' age vary from 18 to 55 years (average 31.6 years) and were of 6 different nationalities (Chinese, Dutch, Iranian, Italian, Polish, and Vietnamese). They were seated at a distance of about 0.5 meter from the monitor in a dimmed room. Using a slider bar they indicated the perceived age of the faces. Four different participants estimated the age of each face. The estimated ages by different participants, have been confirmed to be highly consistent (Cronbach's $\alpha > 0.87$). Average estimated age for each face is used as perceived age in this paper.

5 Experimental Results

In this section, we discuss the results of our experiments. First, we will discuss the accuracy of the system when only facial contrast and intensity (appearance) features are used, either individually or taken together. Then, the effect of color and contrast usage for estimating true and perceived ages will be analyzed.

5.1 Color, Contrast and Appearance

In this paper, we propose to combine facial contrast and appearance, and use perceptual color space CIELAB to increase the age estimation accuracy as well as improving the correlation between automatic estimations and perceived ages. However, it is also important to show the discriminative power of facial contrast and appearance, individually. For this purpose, we evaluate the individual and combined use of these features for different color spaces to estimate true age. In this experiment, a two level 10-fold cross-validation scheme is used on the whole set of the UvA-NEMO Smile Database. Each time a test fold is separated, a 9-fold cross-validation is used to train the system and select the regression parameters. There is no subject overlap between folds. The resulting *mean absolute error* (MAE) is given in Table 2.

Table 2. The MAE for true age estimation using contrast, appearance, and combined features

Features	Mean Absolute Error		
	Gray-scale	RGB	CIELAB
Contrast	9.50 ($\pm$8.31)	8.47 ($\pm$7.71)	8.41 ($\pm$8.04)
Appearance	6.12 ($\pm$4.87)	5.82 ($\pm$4.72)	5.81 ($\pm$4.59)
Combination	6.03 ($\pm$4.63)	5.59 ($\pm$4.52)	5.54 ($\pm$4.54)

Results show that enabling the use of color in the estimation of true age, noticeably improves the performance with respect to the use of gray-scale. CIELAB color space provides the most accurate estimations for both contrast, appearance, combined features. RGB performs slightly worse than CIELAB.

It is clear that using only facial contrast is not enough for an accurate age estimation system. In CIELAB color space, the MAE of using contrast features is 8.41 years where the MAE for facial appearance is only 5.81 years. Nevertheless, by combining the contrast and appearance features, the proposed system is able to achieve the best result with an MAE of 5.54. Combined features increase the accuracy approximately 4% and 5% (relative) with respect to appearance for RGB and CIELAB, respectively. However relative improvement for gray-scale is only 1.47%, which shows that gray-scale contrast is not as informative as color contrast.

5.2 Estimating True and Perceived Age

The aim of this study is to enable perceptual cues for age estimation to improve the correlation between automatically estimated and perceived ages. To this end, we analyze the effect of color and contrast usage for estimating true and perceived ages. As described in section 4.2, we collected perceived ages for 84 of 400 subjects. Face images of the remaining 316 subjects are used for training and validation. The regression parameters are selected using a 10-fold cross-validation on the training set. The trained systems are tested on these 84 subjects. Note that the true age information is used for training and validation in this experiment. Using appearance and combined features, the MAE and the correlation of estimations are computed for true and perceived ages. Reported correlation values are the linear correlation coefficients between the estimated ages and the true/perceived ages. Table 3 shows the MAE and the correlation for true and perceived age estimation using different color spaces.

Table 3. The MAE and the correlation for true and perceived age estimation

Features	True MAE Years (Correlation)		Perceived MAE Years (Correlation)	
	Appearance	*Combination*	*Appearance*	*Combination*
Gray-scale	5.98 (0.89)	5.79 (0.91)	6.52 (0.88)	6.11 (0.90)
RGB	5.46 (0.91)	5.23 (0.92)	6.24 (0.89)	5.08 (0.91)
CIELAB	5.38 (0.91)	5.09 (0.93)	6.08 (0.89)	4.09 (0.94)

Results show that enabling contrast features in the estimation of true age, decreases the MAE and increases the correlation by approximately 4% and 2% on average, respectively. For the estimation of perceived age, the combined use of appearance and contrast features, decreases the MAE and increases the correlation by approximately 19% and 3% on average, respectively. When we analyze the improvement rates, it is seen that the use of facial contrast shifts the estimations towards perceived ages. The relative MAE improvement, using the CIELAB color space, for perceived age estimation is approximately 5 times more than the improvement for true age estimation. Also, the use of color for automatic age estimation decreases the MAE noticeably, where the best accuracy is achieved by using combined features in CIELAB color space.

6 Conclusions

In this study, we have introduced the usage of automatically extracted facial contrast features and perceptual color space to improve age estimation and increase the correlation between automatically estimated and perceived ages. The majority of automatic facial age estimation methods focus on the appearance of the face, as the appearance is the most revealing aspect of aging. However, we show that facial contrast improves the estimation accuracy for both true and perceived ages.

Additionally, we evaluate the effect of using different color spaces on age estimation accuracy. Our results show that color-based features perform better than gray-scale. Besides, in our experiments, using perceptual CIELAB color space has provided the highest estimation performance for both true and perceived ages.

References

1. D. M. Burt and D. I. Perrett. Perception of age in adult caucasian male faces: computer graphic manipulation of shape and colour information. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 259(1355):137–143, 1995.
2. H. Dibeklioğlu, A. A. Salah, and T. Gevers. Are you really smiling at me? Spontaneous versus posed enjoyment smiles. In *ECCV*, 2012.
3. H. Dibeklioğlu, A. A. Salah, and T. Gevers. A statistical method for 2-d facial landmarking. *IEEE Trans. on Image Processing*, 21(2):844–858, 2012.
4. B. Fink and P. Matts. The effects of skin colour distribution and topography cues on the perception of female facial age and health. *Journal of the European Academy of Dermatology and Venereology*, 22(4):493–498, 2008.
5. Y. Fu, G. Guo, and T. S. Huang. Age synthesis and estimation via faces: A survey. *IEEE Trans. on PAMI*, 32(11):1955–1976, 2010.
6. X. Geng, K. Smith-Miles, and Z. Zhou. Facial age estimation by nonlinear aging pattern subspace. In *ACM Multimedia*, pages 721–724, 2008.
7. X. Geng, Z.-H. Zhou, Y. Zhang, G. Li, and H. Dai. Learning from facial aging patterns for automatic age estimation. In *ACM Multimedia*, pages 307–316, 2006.
8. D. A. Gunn, H. Rexbye, C. E. Griffiths, P. G. Murray, A. Fereday, S. D. Catt, C. C. Tomlin, B. H. Strongitharm, D. I. Perrett, M. Catt, et al. Why some women look young for their age. *PLoS One*, 4(12):e8021, 2009.
9. A. A. Michelson. *Studies in optics*. Dover Publications, 1995.
10. C. A. Murphy. The role of perception in age estimation. In *Digital Forensics and Cyber Crime*, pages 1–16. Springer, 2012.
11. A. Nkengne, C. Bertin, G. Stamatas, A. Giron, A. Rossi, N. Issachar, and B. Fertil. Influence of facial skin attributes on the perceived age of caucasian women. *Journal of the European Academy of Dermatology and Venereology*, 22(8):982–991, 2008.
12. T. Ojala, M. Pietikainen, and D. Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1):51–59, 1996.
13. T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. on PAMI*, 24(7):971–987, 2002.
14. A. Porcheron, E. Mauger, F. Morizot, and R. Russell. Faces with higher contrast look younger. *Journal of Vision*, 11(11):635–635, 2011.
15. N. Ramanathan, R. Chellappa, and S. Biswas. Computational methods for modeling facial aging: A survey. *Journal of Visual Languages & Computing*, 20(3):131–144, 2009.
16. K. Ricanek and T. Tesafaye. Morph: A longitudinal image database of normal adult age-progression. In *International Conference on Automatic Face and Gesture Recognition*, pages 341–345, 2006.
17. M. Stokes, M. Anderson, S. Chandrasekar, and R. Motta. A standard default color space for the internet-srgb. *Microsoft and Hewlett-Packard Joint Report*, 1996.
18. J. Suo, S. C. Zhu, S. Shan, and X. Chen. A compositional and dynamic model for face aging. *IEEE Trans. on PAMI*, 32(3):385–401, 2010.
19. The FG-NET Aging Database. http://sting.cycollege.ac.cy/∼alanitis/fgnetaging/index.htm, 2002.
20. M. C. Voelkle, N. C. Ebner, U. Lindenberger, and M. Riediger. Let me guess how old you are: Effects of age, gender, and facial expression on perceptions of age. *Psychology and aging*, 27(2):265, 2012.
21. J.-G. Wang, W.-Y. Yau, and H. L. Wang. Age categorization via ecoc with fused gabor and lbp features. In *Workshop on Applications of Computer Vision*, pages 1–6. IEEE, 2009.
22. C. Zhan, W. Li, and P. Ogunbona. Age estimation based on extended non-negative matrix factorization. In *IEEE Int. Workshop on Multimedia Signal Processing*, 2011.