

Illumination Independent Color-Based Face Detection

Jure Kovač, Peter Peer, Franc Solina
University of Ljubljana, Faculty of Computer and Information Science
Tržaška 25, SI-1000 Ljubljana, Slovenia
jure.kovac@link.si, {peter.peer, franc.solina}@fri.uni-lj.si

Abstract

Computer vision is one out of many areas that want to understand the process of human functionality and copy that process with intention to complement human life with intelligent machines. For better human-computer interaction it is necessary for the machine to see people. This can be achieved by employing face detection algorithms, like the one used in the installation “15 Seconds of Fame” [10]. Mentioned installation unites the areas of modern art and technology. Its algorithm is based on skin-color detection. One of the problems this and similar algorithms have to deal with is sensitivity to the illumination conditions under which the input image is captured. Hence illumination sensitivity influences face detection results. This problem is being more or less successfully solved by the use of color compensation and color constancy methods. In this work some of these methods are described, realized and tested. Their basic intention is to eliminate the influence of non-standard illumination from images. Tests that were performed showed that methods apply positive influence on face detection results.

1 Introduction

1.1 Computer vision and face detection

In the last 30 years one of the most interesting areas of research is building machines that would complement human life with the help of artificial intelligence. This area is full of different challenges and one among them is to imitate human vision. Analogically this discipline is called Computer Vision. The basic idea is to discover properties of a 3D world by using only 2D information from a picture. A lot of effort was put into this area of research, which eventually led to progress in the areas such as object recognition, picture understanding and scene reconstruction. This encouraged new researches with goals to enable a computer to see people, recognizes them and interpret their gestures, expressions and behaviour.

If computer would like to see people, it first has to find

out, where the people are. Next, computer has to find out, who this people are and then what they do or what they want from it.

The whole area is interesting for many applications that are useful in many other areas, such as: human-computer interaction, security and surveillance, entertainment etc.

Searching for faces in 2D images represents a base for realization of previously mentioned ideas. Despite the fact that a lot of research groups all around the world employ this problem and are achieving quite remarkable results, a general solution is yet to be found. The problem of computational complexity and efficiency are proportional to each other, which means that complex systems that could assure good results can't operate in real time and less complex solutions cause greater error. For this reason we have to adapt the algorithm development to the area of final application. Based on our application, we can choose between many approaches to face detection, such as: model-based, feature-based, texture-based, color-based etc.

1.2 Installation “15 Seconds of Fame”

The installation “15 Seconds of Fame” [10] is an interactive art installation, which intends to make instant celebrities out of common people by putting their portraits on the museum wall (Fig. 1). The idea was inspired by the quotation of the famous artist Andy Warhol: “In the future everybody will be famous for 15 minutes” and by the pop-art style of his work. Warhol transformed faces from everyday press to pictures by performing so called “color surgery” on them. First he gathered objects that he wanted to expose by separating faces from the background or by choosing specific face elements, then he applied his famous coloring on these objects. In such fashion he transformed famous people like Mao-Tse Toung, Marilyn Monroe, Jackie Kennedy etc. In the installation 15 minutes from Warhol's quotation is diminished to 15 seconds, which makes the installation more dynamic and at the same time it reflects the dynamics of life in the era of exponential technology, science and art development and fugacity of individuals and their “fame” in today's world.

The installation looks like a valuable framed picture, hanging on the wall (Fig. 1). LCD monitor and digital



Figure 1. LCD computer monitor dressed up like a precious painting. The round opening above the picture is for the digital camera lens.

camera are built into the picture. Camera is connected to a computer, which controls the camera and processes captured images.

Special software contains algorithm for face detection, which looks for faces in captured images. Among them it chooses one for further processing. In the next step a randomly chosen portrait is processed with randomly chosen filter and random coloring is applied afterwards. The random filter is a mixture of commonly used filter types like posterize, color balance and hue-saturation balance. In such a way the portrait in a pop-art fashion arises, which is afterwards shown on screen for 15 seconds, while another image is already being processed.

The algorithm for face detection is based on color information. First it finds skin color pixels in an image, then skin elements are segmented into regions. Unsuitable regions are then eliminated on the basis of geometric properties of the face. Remaining regions represent faces. As illumination conditions affect the colors and this algorithm basis merely on colors, the results of face detection are influenced by illumination. We obtained the best results with our own skin color model, though we also tested other color models [6, 8, 9].

1.3 Illumination problem — Motivation

The purpose of studying methods for eliminating the influence of non-standard illumination in our project is to improve the results of face detection algorithm used in the installation “15 Seconds of Fame” [10]. Non-standard illuminations are by definition those that are more or less different from daylight illumination (defined by CIE standard for illuminant D65) [1]. We find such illumination almost anywhere in enclosed spaces with artificial illumination, where the installation could potentially be exhibited. There are two main groups of methods for addressing this problem: color compensation methods and color constancy methods.

1.4 Structure of the paper

In the next section the methods for eliminating the influence of non-standard illumination are presented. Section 3 compares the methods effectiveness. Method selection for the installation “15 Seconds of Fame” is described in Section 4. We conclude the paper in Section 5.

2 Methods for eliminating the influence of non-standard illumination

2.1 Color compensation methods

Methods in this group have low time complexity (order of $O(n)$) and they do not need a preliminary learning step. This means that they are easy and straightforward to implement. Their effectiveness is relatively high on sets of images with some input constraints. Illumination should be relatively close to standard illumination. The input image is transformed in the way that the colors in the image are leveled in respect to some statistical quantity.

Grey World (GW) [5] algorithm presents simple and fast method for color compensation on images which are defined in *RGB* color space [4]. It is based on the presumption that the average surface color on the image is acromatic. This means that the average color, which is reflected from the surfaces, corresponds to the color of the illumination. To execute the algorithm we have to calculate the averages for each channel *R*, *G* and *B* for the whole image. Averages are then transformed with the help of a linear scale factor to values that correspond to the mean gray value of one standard illuminant.

Modified Grey World [3] method is very similar to basic GW algorithm with the difference that the scale factors are calculated in a slightly different way. The average of each channel is calculated by counting each color only once, regardless of how many times the color appears in the image. By doing so, we eliminate the influence of colors represented by a large number of pixels in the image on the average color. The method is effective on images, which do not contain a lot of different colors.

White-Patch Retinex [5] method is like the Modified GW method just a special version of the basic GW method. The difference lies again in the method of calculating the scaling factors. Instead of the average color we use the maximal value of each channel in the whole image. The Retinex method is above all suitable for dark images.

2.2 Color constancy methods

Methods belonging to this group differ from the color compensation methods above all in the need to integrate a

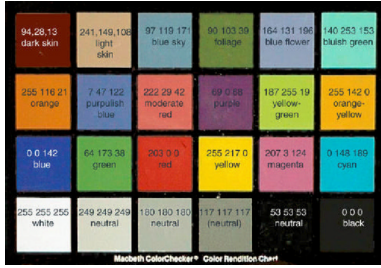


Figure 2. The Macbeth color checker [2].

preliminary learning step. They need the knowledge about illumination properties and properties of the capturing devices, e.g. digital cameras. The input image is then transformed in such a way that it reflects the state, which is independent of the illumination. Thus a stable presentation of colors under different illuminations is achieved. Generally speaking the methods consist of two distinct steps: scene illumination detection and standard illumination reconstruction. In the first step, the algorithm determines with the help of preliminary knowledge which illumination out of the set of known illuminations is present in the image. In the second step, it applies the necessary transformations to reconstruct the standard (or other wanted) illumination.

In the *Color by Correlation* method [3] the preliminary knowledge about the illumination properties is represented with a set of colors, which can appear under specific illumination, i.e. colors that are visible under specific illuminant.

2.2.1 Illumination reconstruction

After the illumination detection based on correlation technique takes place, we need to reconstruct the image scene under some standard illumination conditions. In order to perform such reconstruction, certain transformations should be applied. To calculate the transformation parameters, we need the information about the spectral power distribution. We can gain this information with the help of the Macbeth color checker (Fig. 2) [2]. We need two images of the Macbeth color checker captured under different illuminations. The first one should be captured under the same illumination as the input image, which we want to reconstruct. The second image should be captured under standard (or other wanted) illumination, which we want to use in the reconstruction process.

3 Comparison of methods

3.1 Color compensation methods

In order to determine the influence of these algorithms on our face detection results, some experiments were per-

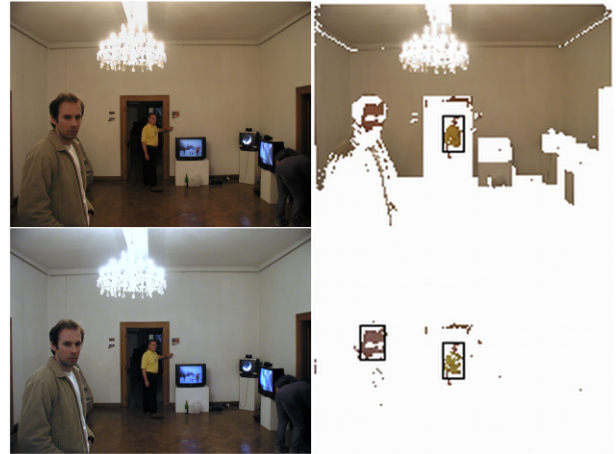


Figure 3. GW performance: the upper part of the figure shows how face detector failed to detect faces due to confusion caused by surrounding colors. This was improved by GW preprocessing as seen in the lower part of the picture. The image was chosen from the incandescent subset of images.

formed on the set of images gathered in our lab and at the first public showing of the installation. The testing set is composed of 160 images taken under four different types of illumination conditions. One subset of images (40 images) was taken under standard daylight, in the second subset (40 images) objects were illuminated by incandescent lamps assembled into a chandelier, the third subset (40 images) was taken under camera flash light conditions, and the last subset of images (40 images) was taken under neon light illumination conditions. After that, one of the color compensation methods was applied and finally, face detection algorithm was applied to original and preprocessed images.

Results gathered in Tab. 1 show perceivable improvement in face detection on images taken under different than standard illumination conditions when one of the compensation algorithms was previously applied (see Fig. 3). Note that the original face detection algorithm was developed to work best under standard daylight illumination [9]. Grey World algorithm performed especially well since for the *flashlight*, *incandescent* and *neon* light conditions a considerable increase in TP/Det percentage can be noticed.

All these results are dependent on our skin detection technique used for face detection, which works in the 3D color space (RGB). Some research suggests that skin detection in 2D color space (YUV) might be better since it is less brightness dependent than detection in the 3D color space (RGB) [4]. Our experiments on the data collected with this installation, however, do not prove the advantage

Method	None	GW	MGW	RET	None	GW	MGW	RET
Illuminant	<i>standard</i>				<i>incandescent</i>			
All Faces	109				95			
Detected	75	70	65	76	45	57	42	43
TP	68	65	60	68	28	45	31	29
FP	7	5	5	8	17	12	11	14
FN	40	44	48	40	67	50	64	66
TP/Det	90,66	92,85	92,31	89,47	62,22	78,95	73,81	67,44
FN/All	36,70	40,37	44,04	36,70	70,53	52,63	67,37	69,47
Illuminant	<i>flashlight</i>				<i>neon</i>			
All Faces	112				78			
Detected	55	47	43	39	63	64	29	60
TP	38	39	36	32	50	54	26	49
FP	17	8	7	7	13	10	3	11
FN	74	73	76	79	28	24	52	29
TP/Det	69,09	82,98	83,72	82,05	77,77	84,37	89,66	81,66
FN/All	66,07	65,18	67,86	70,53	35,90	30,77	66,66	37,18

Table 1. Color compensation results show the number of all detections (Detected), the number of detected faces as true positives (TP), number of false detections as false positives (FP) and number of faces missed as false negatives (FN) on four subsets of images which represent different illumination conditions (standard, incandescent, flashlight and neon), previously preprocessed by Grey World (GW), Modified GW (MGW), White Patch Retinex (RET) or no preprocessing at all (None). Row All Faces shows the number of faces in particular subset of images. TP/Det shows the percentage of true positives out of all detections and FN/All shows the percentage of false negatives out of all faces in the subset. For the installation the first percentage is extremely important, while the second one is merely informative, since we have consciously eliminated faces that were too small for further processing but were included in the number of all faces!

of the 2D color space [8]. Other approaches for adapting skin color model to illumination changes could be used, e.g. [11], but the question with these approaches is whether preserved skin color information would satisfy our segmentation scheme.

The results also show that the performance of these techniques depends very much on the type of the illumination. Therefore a considerable amount of precaution should be taken in decisions about the usage of these techniques. On the other hand all of these algorithms are very effective from the time complexity point of view and as such they enable the possibility of performing a simple initialization test when the scene is first placed under certain illumination. In this way we can determine which algorithm would produce the best results under certain type of illumination.

3.2 Color constancy methods

Tests for the correlation method were performed on a testing set containing approximately 100 images with one or multiple faces, taken under 5 various illumination conditions. The *white* subset represents the standard illumination conditions, while *red*, *blue*, *green* and *yellow* subsets were

generated by illuminating the scene with the corresponding color light.

First, illumination detection on images was performed with correlation technique, then the images were reconstructed under approximation of standard illumination and finally face detection algorithm was applied on images (Figs. 4 and 5).

The results are summarized in Tab. 2. Results of subset *white* are shown as a comparable reference to other results. Columns with no method previously applied (None) can contain zero detections. This can occur if face detection algorithm finds no skin color in an image, which can often be the case in extreme illumination conditions.

Results in Tab. 2 show the positive effect of color correction on images with non-standard illumination conditions. If we do not apply preprocessing, the face detector finds some faces only on images from the yellow subset. After the preprocessing step almost all faces were recovered under this illumination, while under all other illuminations the number of recovered faces was not that high, but the difference with the results gained without preprocessing is enormous.

In the blue subset we see a very large number of false

Illuminant	<i>white</i>	<i>yellow</i>		<i>green</i>		<i>blue</i>		<i>red</i>	
Method	None	None	C	None	C	None	C	None	C
All Faces	38	35		23		42		34	
Detected	43	13	40	0	18	19	80	10	25
TP	35	7	34	0	13	0	29	0	23
FP	8	6	6	0	5	19	51	10	2
FN	2	28	1	23	10	42	13	34	11
TP/Det	81,39	53,85	85,00	0	72,22	0	36,25	0	92,00
FN/All	5,26	80,00	2,86	100,00	43,48	100,00	30,95	100,00	32,35

Table 2. Correlation results show the number of all detections (Detected), number of correct face detections as true positives (TP), number of detections that turned out not to be faces as false positives (FP) and the number of faces missed by detection algorithm as false negatives (FN) for different subsets of images (white, yellow, green, blue and red), previously processed by correlation algorithm (C) and with no preprocessing at all (None). Row All Faces shows the number of faces in a particular subset of images. TP/Det shows the percentage of true positives out of all detections and FN/All shows the percentage of false negatives out of all faces in the subset. The TP/Det is for the installation extremely important, while FN/All is merely informative for the performance of our face detector. Note that small faces are deliberately eliminated from further processing already by the face detection algorithm.



Figure 4. Correlation performance on image from the yellow subset: detection results on image without preprocessing (top) and after illumination reconstruction (bottom).

detections caused by the mixture of incandescent and blue light. This mixture was necessary for enhancement of other than blue color channels (R and G) since blue filter was very strong. If only blue channel is present, we have as much information as in achromatic pictures. Incandescent light caused an interesting effect of appearing all shadows slightly yellow in reconstructed pictures and as a consequence many false faces were found.

Nevertheless, with the exception of blue illumination these results are quite comparable with the results in Section

Figure 5. Correlation performance on images from the green (left) and the blue (right) subset. On the original image (top) no faces could be detected. After correcting the illumination (middle) three faces were detected.

3.1, where face detection was tested under close to standard illumination conditions.

4 Method selection

The nature of illumination in some galleries can represent a real problem since it normally differs from daylight illumination. In older rooms in particular we usually have chandeliers or similar types of illumination that emit a prevailing yellow light into the room and as a consequence a prevailing yellow color in captured images. This causes the shift of a large part of image color space into the color subspace, which represents skin color. An illustrative example of this property can be observed on the white walls of a room. Normally they are white, but under incandescent lamp illumination they are more bright yellow than white. And since walls can occupy large parts of an image, it can happen that most of the image pixels are recognized as skin-like pixels (see Figs. 3 and 4). This type of illumination can have a serious negative influence on the number of false face detections (false positives).

In case of incandescent lamp illumination we should choose among color compensation methods described in Section 2.1. Based on the results of these algorithms and constraints discussed in Section 3.1, we decided to use GW algorithm as it performs best when minor lighting deviations from standard illumination are present. Although, some form of automatic selection is taken into future consideration.

A totally different story can be observed in discotheques, where illumination emits color light (e.g. blue, green, red etc.) into the room. This shifts all scene colors towards the color of the illumination. Consequently, a lot of skin-like pixels are recognized as non-skin-like pixels and the number of correctly detected faces (true positives) is decreased, since we can not reliably find skin-like pixels.

When deviations from standard illumination are much more noticeable, we must choose a correlation technique with proper illumination reconstruction.

Eliminating the influence of non-standard illumination before face detection ensures much better results. The whole system is much more flexible and the installation can be exhibited almost anywhere.

5 Conclusion

The problem of elimination of non-standard illumination is one of the most burning problems in the area of computer vision. Most today techniques do not provide satisfying results, which would be very much desired and sometimes even urgent. Based on that fact it is sometimes even necessary to replace methods that are influenced by illumination conditions (color based) by other methods, which are more illumination independent (geometric based) and hence more computationally pretentious.

In this work some methods for eliminating the effect of illumination are realized and tested. These methods, despite of all mentioned problems and constraints, serve the purposes of the final application – installation “15 Seconds of Fame”. On the other hand, we have to note that illumination conditions should not be very extreme. With mild deviation of illumination conditions from standard daylight illumination [1], results shown in Tabs. 1 and 2 expose considerable improvement of face detection results. We can conclude that the primary goal of eliminating the illumination condition constraint from face detection algorithm has been achieved.

References

- [1] Joint ISO/CIE Standard ISO 10526:1999/CIE S005/E-1998 CIE Standard Illuminants for Colorimetry. 1999. <http://www.cie.co.at/cie/>
- [2] G. D. Finlayson. Color Constancy in Diagonal Chromaticity Space. In *Proceedings of the 5th IEEE International Conference on Computer Vision (ICCV'95)*, pages 218–223, Cambridge, MA, USA, 1995.
- [3] G. D. Finlayson, S. D. Hordley, and P. M. Hubel. Color by Corellation: A Simple, Unifying Framework for Color Constancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1209–1221, 2001.
- [4] A. Ford and A. Roberts. Colour Space Conversions. 1998. <http://www.inforamp.net/~poynton/PDFs/coloureq.pdf>
- [5] B. Funt, K. Bernard, and L. Martin. Is Machine Colour Constancy Good Enough. In *Proceedings of the 5th European Conference on Computer Vision (ECCV'98)*, pages 445–459, Freiburg, Germany, 1998.
- [6] R. L. Hsu, M. Abdel-Mottaleb, and A. K. Jain. Face Detection in Colour Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):696–706, 2002.
- [7] J. Kovač. Eliminating the Influence of Non-Standard Illumination from Images. *Diploma Thesis*, Faculty of Computer and Information Science, University of Ljubljana, Ljubljana, Slovenia, 2002.
- [8] J. Kovač, P. Peer, and F. Solina. 2D Versus 3D Color Space Face Detection. *Technical report*, Computer Vision Laboratory, Faculty of Computer and Information Science, University of Ljubljana, 2003. <http://eprints.fri.uni-lj.si>
- [9] P. Peer and F. Solina. An Automatic Human Face Detection Method. In *Proceedings of the 4th Computer Vision Winter Workshop (CVWW'99)*, pages 122–130, Rastendorf, Austria, 1999.
- [10] F. Solina, P. Peer, B. Batagelj, and S. Juvan. 15 Seconds of Fame – an Interactive, Computer-Vision Based Art Installation. In *Proceedings of the 7th International Conference on Control, Automation, Robotics and Vision (ICARCV'02)*, pages 198–204, Singapore, 2002.
- [11] M. Störring and E. Granum. Adapting a Statistical Skin Colour Model to Illumination Changes. In *Proceedings of the 1st European Conference on Color in Graphics, Imaging and Vision (CGIV'02)*, pages 16–21, Poitiers, France, 2002.