Eliminating the Influence of Non-Standard Illumination from Images

Jure Kovač, Peter Peer, Franc Solina

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

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" [8]. 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.

Keywords

computer vision, automatic detection, human face, face candidates search, illumination reconstruction, color compensation, color constancy

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 inteligence. 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 jet 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" [8] 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 world famous for fifteen 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 camera are built into the picture. Camera is connected to a computer, which controls the camera and processes captured images.

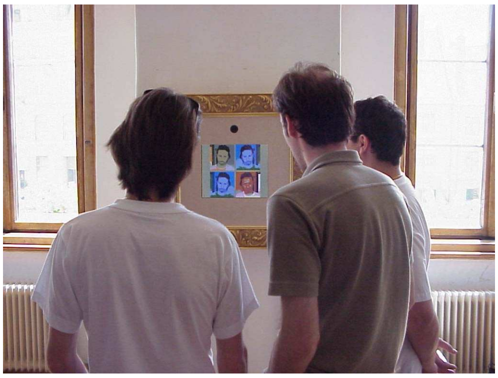

a)

b)

Figure 1: Installation at Display: a) LCD computer monitor dressed up like a precious painting. The round opening above the picture is for the digital camera lens. b) Group of people in front of the installation.

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.

Additional rules have been applied for variegation of action. In 75% of cases one large portrait is shown by the installation and in 25% of cases four diminished portraits are shown. Each of the four smaller portraits can be processed with a different ?lter and the right-hand portrait column can be mirrored along the vertical axis. This way of stacking together multiple images also resembles Andy Warhols way of displaying and amassing of images. Installation also does not show the same person in sequential intervals. If no face is found, the last found face is being processed with different filter. The counter which counts down the time of one's fame and ID number for ordering the portrait over the internet are also added.

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.

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" [8]. Non-standard illuminations are by definition those that are more or less different from daylight illumination (defined by CIE standard for illuminant D65) [5]. 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 effectivness. 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.

2.1.1 Grey World

Grey World (GW) [4] algorithm presents simple and fast method for color compensation on images which are defined in RGB color space [3]. 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. The corresponding scale factor for each channel R, G and B is calculated as

$$S_C = \frac{C_{std}}{C_{ave}}, \tag{1}$$

where C is one of the channels R, G and B, std stands for standard gray value and ave stands for average value of the channel in the whole image. We change the input image so that for each pixel we multiply the pixel's R, G and B values with corresponding scale factor (Eq. 1).

There are several possible values that can represent the mean gray value. One example of such a value is $\frac{1}{2}RGB_{std}$, which represents one half of ideal gray color under standard illumination (CIE D65 [5]). On the other hand, we can choose a similar constant value, which is reflected by the previously calculated average or by some test results.

The method is very effective if we have to correct the illumination in images that were captured in conditions close to standard illumination. Though the method is powerless in case of extreme illumination conditions, e.g. lights in a discotheque. The root of the problem can be found in the already mentioned presumption that the color of the illumination is equal to the average color of the image. Such a presumption is of course naive for realistic color reconstruction.

2.1.2 Modified Grey World

Modified GW [2] 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. In the basic GW method prevailing colors can have big influence on the average color, resulting in lowering the influence of other colors in the image. The modified GW ensures that all the colors in the image are equally important in calculations.

For calculating the average we need information about all the colors in the image. This information can be gathered with the help of color histograms. This brings additional processing in the algorithm, making it more computationally demanding than the basic GW method.

2.1.3 White-Patch Retinex

Retinex [4] 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. In case of Retinex, instead of the average color we use the maximal value of each channel in the whole image (compare with Eq. 1)

$$S_C = \frac{C_{std}}{C_{\max}}. \tag{2}$$

The Retinex method is above all suitable for dark images. In intensively illuminated scenes the maximal value of each channel is close to the saturation value of 255. Necessary changes in such cases do not take place. The complexity of the Retinex method is the same as the complexity of the basic GW method.

2.2 Color constancy methods

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

2.2.1 Color by Correlation

In the Color by Correlation method [2] 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.

Processing is done in 2D color space [3] and it enables us to reconstruct the color of illumination up to the precision of a multiplicative constant. With this estimation we are able to transform colors within the input image into an illumination independent state. After this illumination independent state is known (image illuminant is detected), it is possible to reconstruct the image under other possible illuminations, e.g. standard illumination.

For 2D color space YUV space was chosen, because its usage is widely spread and it is similar to the TSL color space in the sense of chroma and brightness separation. TSL color space proved itself very well regarding face detection [9]. U and V coordinates represent chroma and Y represents brightness caused by illumination.

The Color by Correlation method consists of three steps. First, we have to code the illumination properties that represent the preliminary knowledge about which colors are possible under specific types of illumination. This knowledge is in the second step correlated to the color information in the actual input image. For each possible illumination we calculate the possibility that this illumination represents the illumination used in the input image. In the third step, we use this result to reconstruct the required illumination.

Correlation matrix $\mathbf{M}$ (Tab. 1) is used to code the preliminary knowledge. First of all, we have to define the range of colors for each possible illumination, which are visible under specific illumination. We can obtain this information by taking one or more pictures of differently colored objects (e.g. Macbeth color checker in Fig. 2) under the same illumination and check which colors are present in the image. With the help of this information we compose the probability distribution. This distribution presents the probability of appearance of each color under specific illumination. Probability distributions for all possible illuminations compose the columns of correlation matrix $\mathbf{M}$ (Tab. 1).

As we compose the correlation matrix $\mathbf{M}$, we compose also the vector $\vec{\mathbf{v}}$ out of the input image by counting pixels of the same color (N_{x_i,y_j}). The vector components are set to 1 or 0 according to Eq. 3

$$\vec{\mathbf{v}}_{x_i,y_j} = \begin{cases} 1 & ; \quad N_{x_i,y_j} > 0 \\ 0 & ; \quad \text{otherwise.} \end{cases} \tag{3}$$

If in vector $\vec{\mathbf{v}}$ there is a 0 for color (x_i, y_j), then this color is not represented in the image.

For each color in the image we calculate the probability P_c. The probability is calculated by counting all pixels of the same color (N_c or N_{x_i,y_j}) and

color	ill_1	ill_2	ill_3	$\ldots$	$\ldots$
$x_1,\, y_1$	$p(x_1, y_1 \vert ill_1)$	$p(x_1, y_1 \vert ill_2)$	$p(x_1, y_1 \vert ill_3)$	$\ldots$	$\ldots$
$x_1,\, y_2$	$p(x_1, y_2 \vert ill_1)$	$p(x_1, y_2 \vert ill_2)$	$p(x_1, y_2 \vert ill_3)$	$\ldots$	$\ldots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_1,\, y_n$	$p(x_1, y_n \vert ill_1)$	$p(x_1, y_n \vert ill_2)$	$p(x_1, y_n \vert ill_3)$	$\ldots$	$\ldots$
$x_2,\, y_1$	$p(x_2, y_1 \vert ill_1)$	$p(x_2, y_1 \vert ill_2)$	$p(x_2, y_1 \vert ill_3)$	$\ldots$	$\ldots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$x_n,\, y_n$	$p(x_n, y_n \vert ill_1)$	$p(x_n, y_n \vert ill_2)$	$p(x_n, y_n \vert ill_3)$	$\ldots$	$\ldots$

Table 1: Correlation matrix $\mathbf{M}$: The first row and the first column are not part of the matrix, their purpose is merely representational. In the first column colors in 2D color space [3] are given. Matrix columns represent probability distribution of colors for each possible illumination. Each cell in the matrix reveals the probability of color $x_i,\, y_j$ under illumination ill_k.

then dividing that number with the number of all pixels in the image (N_{all})

$$P_c = \frac{N_c}{N_{all}}. \tag{4}$$

The correlated value of two input images tells us how the input images coincide. The bigger is their correlation value the more alike are the input images. For instance; vectors $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ highly coincide if the value $\vec{\mathbf{a}} \circ \vec{\mathbf{b}}$ is big. The correlation value is computed by operation $\circ$ (correlation)

$$\vec{\mathbf{a}} \circ \vec{\mathbf{b}} = \vec{\mathbf{a}}^{\mathrm{T}} \cdot \vec{\mathbf{b}}. \tag{5}$$

Now we have to determine the correlation between each of the possible illuminations and the input image. This can be done by calculating the vector

$$\vec{\mathbf{l}} = \vec{\mathbf{v}}^{\mathrm{T}} \circ \mathbf{M}. \tag{6}$$

The elements of this vector present the correlation value for each of possible illuminations.

In the final step, we have to estimate the color of the actual illumination by finding the illumination, which has the maximal correlation with the input image

$$\begin{aligned} \hat{\mathbf{c}} &= Max(\vec{\mathbf{l}}) \cdot \mathbf{C}_{ill} \\ Max(\vec{\mathbf{x}}) &= \vec{\mathbf{x}}'; \quad \vec{\mathbf{x}}'_i = \begin{cases} 1; & \vec{\mathbf{x}}_i = \max(\vec{\mathbf{x}}) \\ 0; & \text{otherwise,} \end{cases} \end{aligned} \tag{7}$$

where $\hat{\mathbf{c}}$ is a 2-component (U,V) color vector, which represents the estimation of color of illumination, $\mathbf{C}_{ill}$ is a matrix with dimensions $N \times 2$ that stores the estimations of colors of N illuminations, between which we choose from, and function Max returns a vector in which a value 1 is associated only with the component with the maximal value of input vector, while all other components are set to 0. $\vec{\mathbf{l}}$ is expressed in Eq. (6). $\hat{\mathbf{c}}$ actually represents the detected illumination and $\mathbf{C}_{ill}$ represents all possible illuminations!

The correlation method solves the problem of color constancy using the maximal probability approach. Within such a correlation framework we could also implement other algorithms [2], e.g. Gray World, neural networks based approach, Convex Gamut Mapping, Color by Voting etc. All we need to do is to change the correlation matrix in a way that it reflects the presumptions, constraints and connections between colors and illuminations suggested by the mentioned algorithms.

2.2.2 Illumination reconstruction

After the illumination detection based on correlation technique takes place (value $\hat{\mathbf{c}}$ in Eq. 7), we need to reconstruct the image scene under some standard illumination conditions. In order to perform such reconstruction, certain transformations should be applied. In our work this transformations are linear and based on straight line calculations.

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) [1]. 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.

The Macbeth color checker is a device used in professional photography, where it serves as a reference for determining lighting conditions. The checker consist of 24 boxes that are of different colors. The chosen colors represent natural objects like human skin, plants, sky etc. Its purpose is mainly to help in the process of recognition and reconstruction of listed objects under different illuminations.

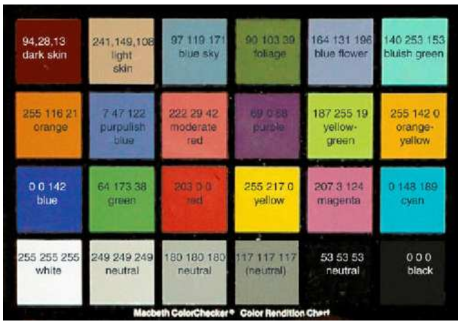


Figure 2: The Macbeth color checker [1].

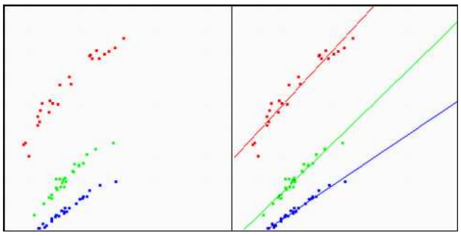


Figure 3: Graph of Macbeth's color averages for standard and unknown illumination: On the left side are shown the points of average values of each color box for all three color channels R, G and B. On the right side the linear approximation is given. x axis denotes the standard illumination averages and y axis represents the unknown illumination averages.

We obtain the transformation parameters by calculating the average values of all color channels R, G and B in each color box of Macbeth's checker. We do that for both checkers, i.e. for the one captured under standard illumination and for the one captured under unknown illumination. These values can be presented on a graph where we have standard illumination averages on the x axis and unknown illumination averages on the y axis (Fig. 3). We can observe that the points that belong to the same color channel indicate a linear shape, which can be well approximated with a straight line using the min-square method. The results of the min-square method reveal the parameters k and n of the straight line model $y = kx + n$ for each color channel (Fig. 3). This straight line model is then used to transform the colors from one illumination into the colors of the second (target) illumination. In this way, the transformation is simple and fast.

We can put these transformations for all possible illuminations also directly into the matrix $\mathbf{C}_{ill}$ (Eq. 7), if we want that vector $\hat{\mathbf{c}}$ carries the information about the needed transformation.

3 Comparison of methods

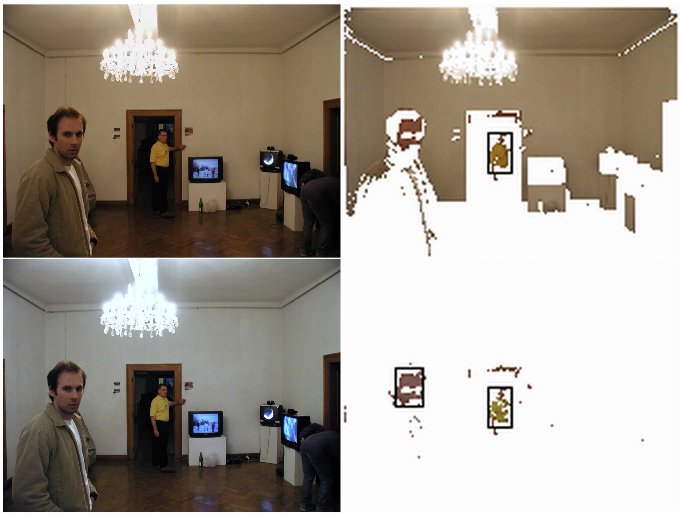


Figure 4: GW performance: the upper part of the figure shows how face detector failed to detect faces due to confusion caused by surrounding colors. This was improved by GW preprocessing as seen in the lower part of the picture. The image was chosen from the incandescent subset of images.

3.1 Color compensation methods

In order to determine the influence of these algorithms on our face detection results, some experiments were performed 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. 2 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 Figs. 4 and 5). Note that the original face detection algorithm was developed to work best under standard daylight illumination [7]. Grey World algorithm

Method	**None**	**GW**	**MGW**	**RET**	**None**	**GW**	**MGW**	**RET**
Illuminant	*standard*				*incandescent*			
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	*flashlight*				*neon*			
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 2: 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 preprocession 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!

Figure 5: Retinex performance: In the upper part of the figure we see two false detections, while on the lower part this problem was overcome by the Retinex algorithm. The image was taken from the flashlight subset.

performed especially well since for the *flashlight, incandescent* and *neon* light conditions a considerable increase in TP/Det percentage can be noticed. Whenever another algorithm performed better than Grey World, this had a significant influence on the FN/All percentage. Which means that the advantage has been gained not by increasing of true positives but by decreasing of false positives and as a consequence false negatives increased. We can easily observe this behavior in the Modified Grey World algorithm results. Either way, if the number of true positives is not too small, this has a very positive effect on our application, since from the application point of view it is crucial not to show false detections too often. A measure againts displaying false positive face detections is also built in the selection mechanism, which in principle selects faces at random, but still gives a higher priority to those faces which are bigger and higher in the input image.

All preprocessing techniques showed little or no improvements at standard illumination conditions. This was somehow expected since the original face detection algorithm was developed under presumption that standard illumination is present in the image.

All these results are dependent on our skin detection technique used in face detection algorithm, where skin color is detected which works in the 3D color space (RGB). Skin detection in 2D color space (YUV) might improve this results considerably as it is less brightness dependent than detection in the 3D color space (RGB) [3].

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: Illumination detection

Tests for the correlation method were performed in order to determine the best merit (function) for illumination detection. The testing set contained 156 images with none, one or multiple faces, taken under 7 various illumination conditions.

The *white* subset contains images taken without any light except for the flash light of the camera. This set represents the standard illumination conditions as it approximates conditions of standard daylight most closely. The color of other illumination conditions can be recognized from the name. Subset red_1 and $blue_1$ are taken under very extreme lighting conditions of particular color, while *red*, *blue*, *green* and *yellow* subsets represents more mild illumination conditions. Illuminations conditions were created with one regular incandescent 40 Watt lamp and one 40 Watt red lamp. Other colors were applied by creating a simple filter out of plastic glasses, which were put over an incandescent lamp.

Correlation merits represent the numbers stored in the correlation matrix and consequently they represent a base for determining the amount of correlation among two image illuminations. P stands for probability of color under particular illumination, $\log(P)$ is the logarithm of that probability, t/f (true/false) merit shows only whether certain color is possible under the illumination in question.

Results (Tab. 3) show that the best illumination detection is achieved by t/f merit, i.e. the smallest number of wrong decisions were made by this merit. We have also to take into consideration the fact that most false detections appeared at extreme illuminations and are due to the lack of color information. This means that one or two color channels at these illuminations had extreme values in almost every image in the subset. Also the majority of these false illumination detections were detected as similar illuminations, e.g. red_1 was detected as *red*.

Good detection results were also achieved by the probability merit, if we don't need detection of standard or close to standard illumination. The rea-

Light	N_{Pics}	$\log(P)$	P	t/f
blue	26	1	0	0
$blue_1$	20	8	0	6
green	18	0	0	0
red	29	10	8	1
red_1	20	20	1	20
yellow	23	0	0	0
white	20	0	20	0
$\sum$	156	29	29	27

Table 3: Performance of correlation method: Illumination detection based on comparison of different correlation merits (probability, logarithm of probability and true/false merit). The table represents numbers of false illumination detections in the selected subset of pictures, except for N_{Pics} column, which stands for the number of pictures in a subset.

son for these false detections lies in probability merit which is biased in favor of colors that are found in other than standard (*white*) illuminations. For instance, blue color will have a very high probability under blue illumination since blue is the predominating color in such illumination, while under standard illumination blue will have a relatively small probability as there will be many other colors present beside it. Under standard illumination more colors are possible and that is why their probabilities are relatively small. As such they are too small to beat other illuminations. Logarithm of probability and t/f merit reduce this bias, but are facing some other types of false detections. These false detections mostly occur at extreme illumination conditions ($red1$, $blue1$) and are often mistaken for the similar illuminations ($red1$ is labeled *red*), but almost never two illuminations of totally different types are mixed up.

3.3 Color constancy methods: Illumination reconstruction

Testing was performed basically on the same set of images as described in section 3.2. Subsets red_1 and $blue_1$ were removed since their illumination is to extreme to be successfully reconstructed, also some images which contain no faces were eliminated from this experiment.

First, illumination detection on images was performed with correlation technique, then the images were reconstructed under approximation of standard illumination (subset *white*) and finally face detection algorithm was applied on images (Figs. 6 and 7).

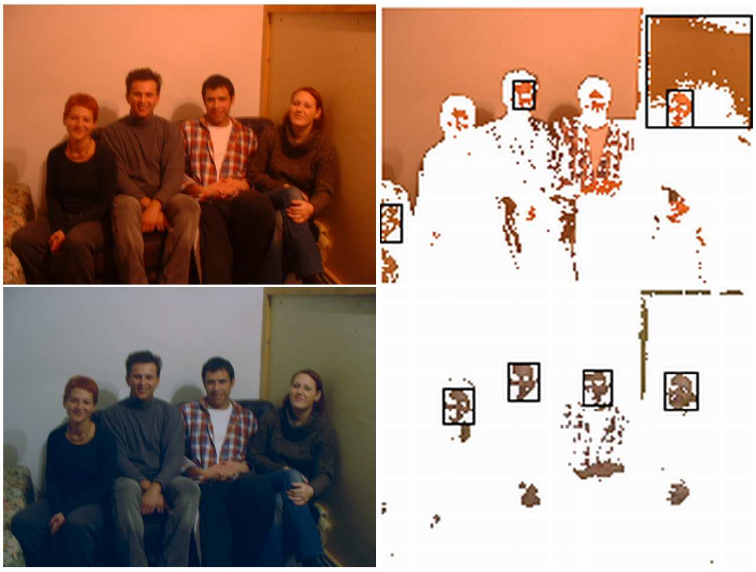


Figure 6: Correlation performance on image from the yellow subset: detection results on image without preprocessing (top) and after illumination reconstruction (bottom).

The results are summarized in Tab. 4. 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. 4 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 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.

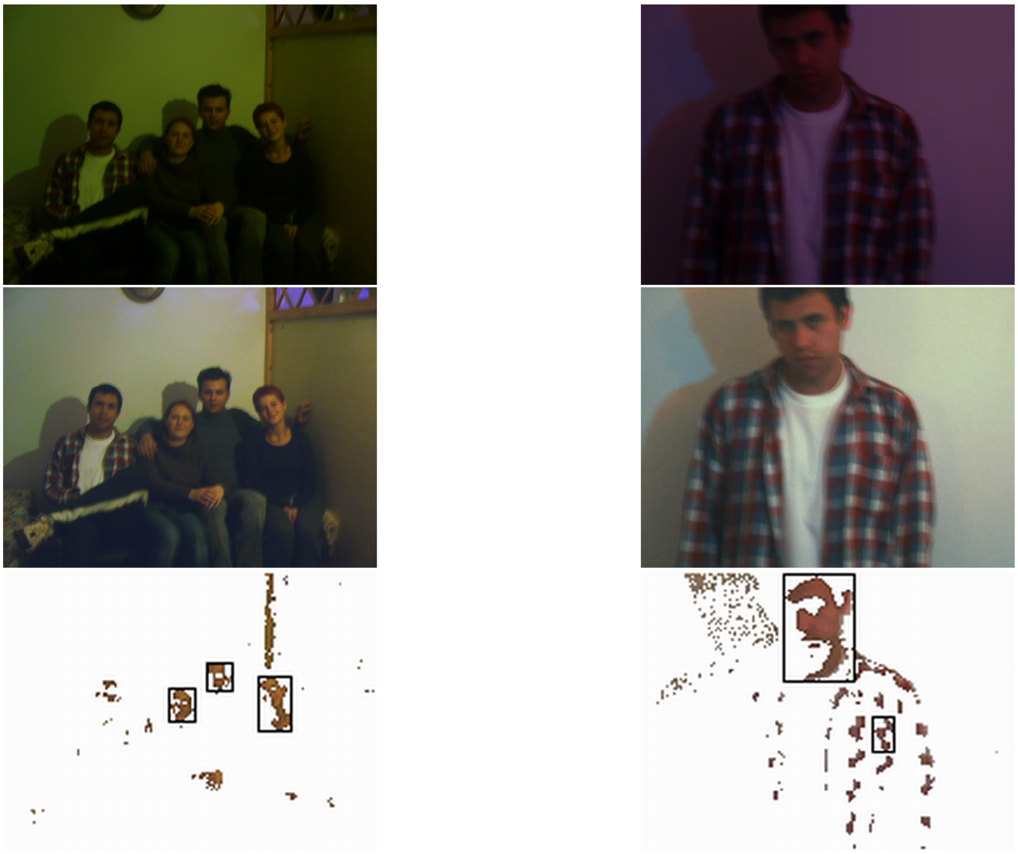


Figure 7: Correlation performance on images from the green (left) and blue (right) subset. On the original image (top) no faces could be detected. After corrected illumination (middle) three faces could be detected.

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

Illuminant	*white*	*yellow*		*green*		*blue*		*red*	
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 4: 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.

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. 4 and 6). 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. When illumination conditions are constant over a large period of time, no illumination detection is necessary. By manually selecting the illumination we eliminate all the risks linked with false illumination detection and assure the best 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”. Some examples of final results of the installation produced by the assistance of described methods are shown in Fig. 8. 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 [5], results shown in Tabs. 2 and 4 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.

Fig. 9 reveals the sequence of steps in the installation “15 Seconds of Fame” after the integration of described step, which eliminates the influence of non-standard illumination.

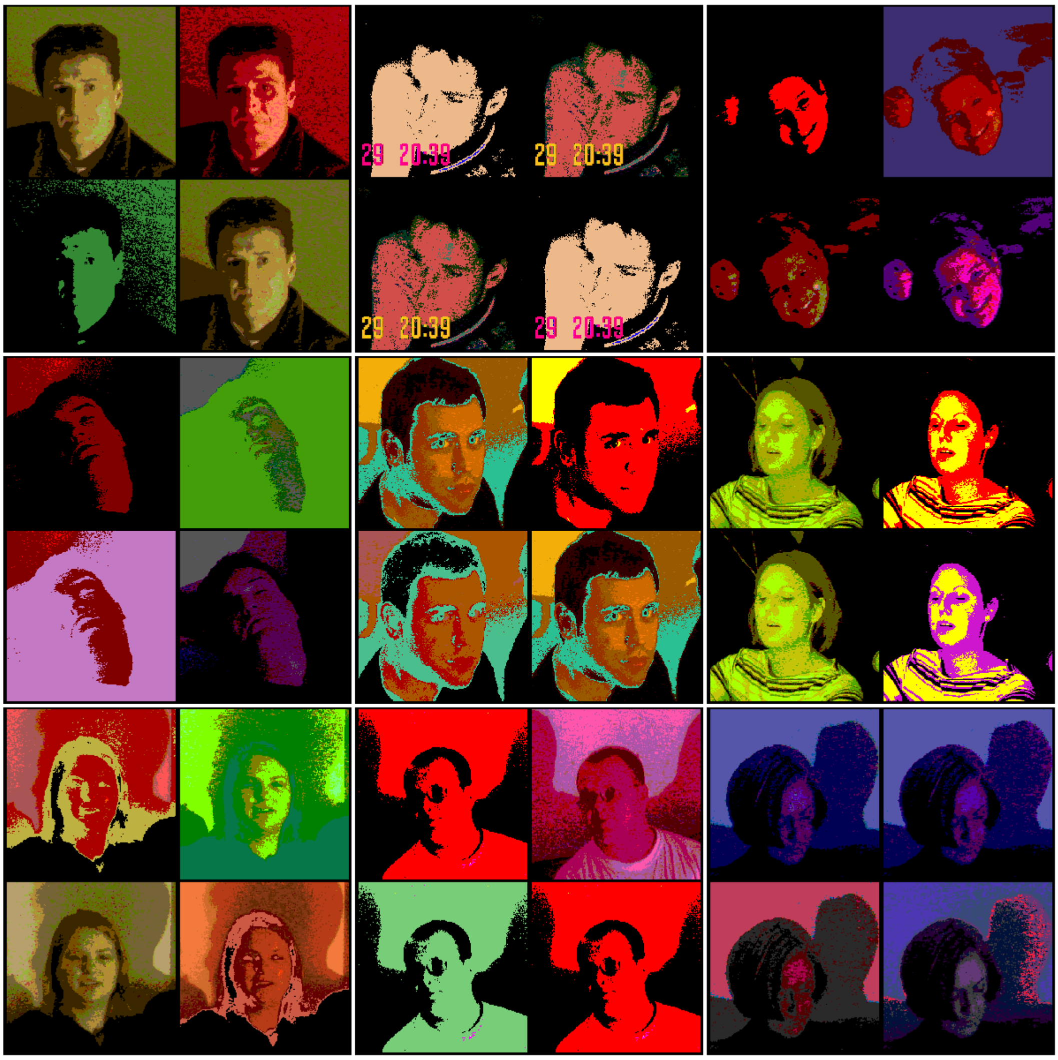


Figure 8: Examples of final results of the installation “15 Seconds of Fame” (pop-art portraits). Pictures were taken from a subset described in Section 3.3

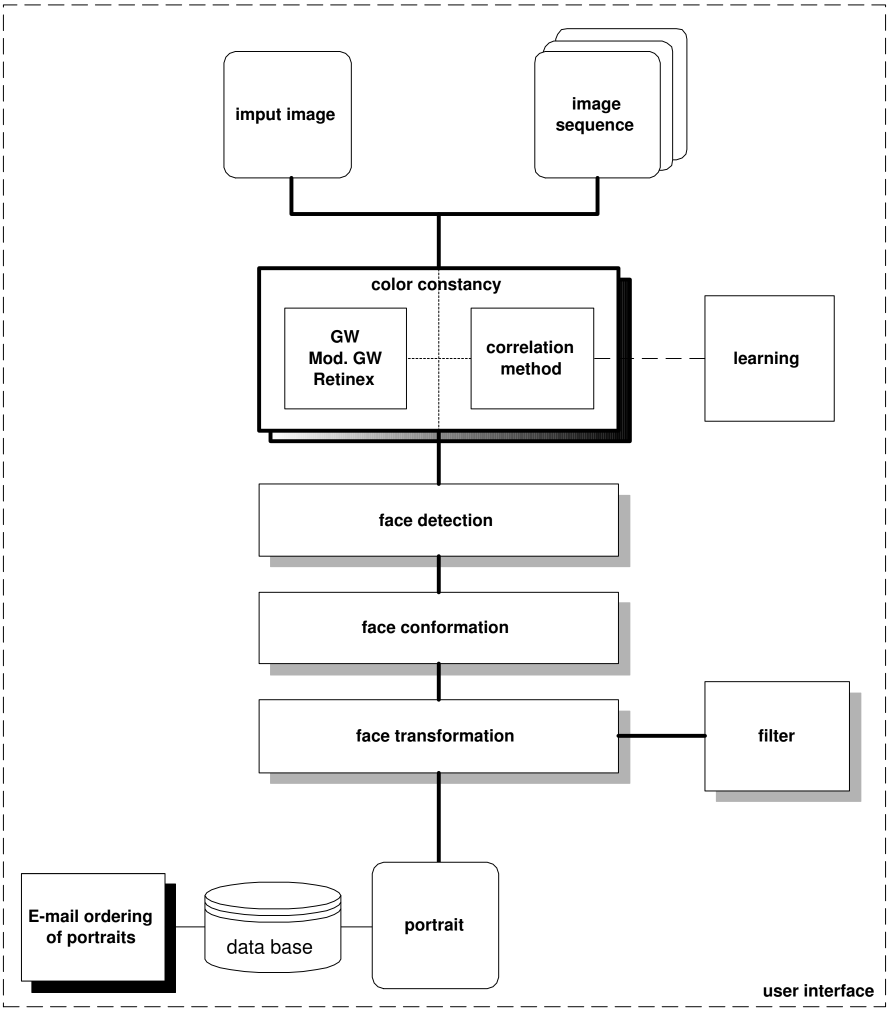


Figure 9: Operational scheme of the instalation "15 Seconds of Fame".

References

[1] G. D. Finlayson, "Color Constancy in Diagonal Cromathicity Space," *Proceedings of the 5th International Conference on Computer Vision (ICCV'95)*, pp. 218–223, 1995.

[2] G. D. Finlayson, S. D. Hordley, P. M. Hubel, "Color by Corellation: A Simple, Unifying Framework for Color Constancy," *IEEE Transactions on Pattern Analysis and Machine Inteligence*, vol. 23, no. 11, pp. 1209–1221, 2001.

[3] A. Ford, A. Roberts, Colour Space Conversions, 1998, *http://www.inforamp.net/˜poynton/PDFs/coloureq.pdf* .

[4] B. Funt, K. Bernard, L. Martin, "Is Machine Colour Constancy Good Enough," *Proceedings of the 5th European Conference on Computer Vision (ECCV'98)*, pp. 445–459, 1998.

[5] Joint ISO/CIE Standard ISO 10526:1999/CIE S005/E-1998 CIE Standard Illuminants for Colorimetry, 1999, *http://www.cie.co.at/cie/* .

[6] J. Kovač, "Eliminating the Influence of Non-Standard Illumination from Images," *Diploma Thesis*, University of Ljubljana, Faculty of Computer and Information Science, 2002.

[7] P. Peer, F. Solina, "An Automatic Human Face Detection Method," *Proceedings of Computer Vision Winter Workshop*, Ed. N. Brändle, pp. 122130, Rastenfeld, Austria, 1999.

[8] Franc Solina, Peter Peer, Borut Batagelj, Samo Juvan, "15 seconds of fame - an interactive, computer-vision based art installation," *Seventh International Conference on Control, Automation, Robotics and Vision ICARCV 2002*, pp. 198–204, Singapore, 2002.

[9] J. C. Terrillon, M. N. Shirazi, H. Fukamachi, S. Akamatsu, "Comparative Performance of Different Skin Chrominance Models and Chrominance Spaces for the Automatic Detection of Human Faces in Color Images," *Proceedings of 4th IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 54–61, 2000.