Interactive Genetic Algorithm for choosing suitable colors in User Interface

Luigi Troiano[1], Cosimo Birtolo[2], and Gennaro Cirillo[2]

[1] Department of Engineering – University of Sannio
RCOST, Viale Traiano – 82100 Benevento, Italy
troiano@unisannio.it

[2] Poste Italiane s.p.a. – Tecnologie dell'Informazione
Centro Sviluppo Servizi Innovativi Napoli
Piazza Matteotti 3 – 80133 Naples, Italy
{birtoloc,ciril127}@posteitaliane.it

Abstract. Designing accessible interfaces does not necessarily mean making them boring, and abandoning an original color scheme. In this paper we present an approach that explores the color palette space, searching for a solution that represents a good compromise between aesthetics and accessibility requirements. This search is performed by using an interactive genetic algorithm which takes into account two different factors made by objective evaluation depending on color distances and contrast ratios and subjective one depending on the score given by user.

1 Introduction

Color vision impaired users perceive colors differently from normal users. This means that although original colors could meet the required luminance contrast ratios for a normal user, the same colors, as perceived by visually impaired users, could not meet the same requirements. Perceived colors can show up with a lower contrast ratio, making difficult for vision impaired audience to access information and services. This requires to adopt color palettes that do not cause significant discomfort to users with color vision deficiencies (CVDs). This does not mean to renounce to the original designer chromatic idea and to make interfaces that are not attractive or boring. It is possible to look for a trade-off between chromatic choices and accessibility for impaired users. This requires to find among the possible color combinations, the palette providing high luminance contrast ratio, but still preserving human preferences that are made by:

– the original chromatic choice of designer
– users' feedback

inal As user-system interaction is the key focus, designing user interface attains to human creativity, perception and aesthetics. For color selection this aspect stands out. Choosing colors is part of the artistic process that leads to outline an user interface.

The problem of finding the right palette for the user interface is a common problem for a designer. Finding an appropriate color palette is an optimization problem with combinatorial complexity as it is necessary to meet requirements such as:

- High luminance contrast between correlated colors (e.g. foreground and background) in order to improve legibility
- Preservation of chromatic choices and preference as expressed by interface designers (e.g. to preserve the meaning of colors)
- To guarantee color accessibility to a broader audience (e.g color vision impaired users).

These requirements can often be conflicting. Considerations above suggest that genetic algorithms are able to solve the problem.

In previous studies [6] the application of genetic algorithms has been tested in order to explore the palette space, and to automatically identify alternatives that can be suggested to the designer. The solution foreseen in this paper can also be used to optimize a color palette for a generic user, when color usability aspects are not taken into account at the beginning of color selection, or when it is necessary to optimize the interface for different usages. Moreover, there are some CVDs, that although well investigated by medicine, does not exist a model able to simulate the deficient perception of colors, in a similar fashion as done by Viénot et al. [2]. In their work they focused attention to how colors are perceived by protanopes and deuteranopes. This is due to the difficulties in building general models for some disorders. For instance, users affected by cataract (a disease making an area of the eye cloudy, so to limit a person from viewing correctly) differs quite a lot from each other, that building a generic model for this disorder is not applicable. When no model exists, the user's preferences become essential in order to make a well-designed and accessible interface.

This paper analyzes how it is possible to overcome this aspect including user evaluation with an Interactive Genetic Algorithm (IGA). In other words, we develop an IGA in order to identify the palette able to maximize the contrast between contiguous colors, and to minimize chromatic differences with designer's scheme, guaranteing color accessibility to impaired users at the same time. For these reasons, the fitness of our algorithm is made by two components: (i) a preference factor depending on some subjective user evaluations of the interface colors in order to capture the user feedback along the evolution process; and (ii) a colorimetric factor as a function of some metrics aimed at evaluating the *contrast ratio* [13] between colors and the distance to the designer's chromatic choice. Unlike to evolutionary algorithms with no feedback, one of the daunting challenges of IGAs entails the effective methods of combating user fatigue. According to [7], user fatigue is a critical element to produce high quality solutions. Evaluating solutions until convergence leads to tedious and demanding attention periods on the user side. We adopted clusters and interpolation as a means for reducing the user fatigue during evolution. The benefit is that user is demanded for evaluating only few individuals and, at the same time, a larger number of solutions can be explored.

The remainder of this paper is organized as follows: Section 2 provides a brief overview of color deficiencies and how these have been addressed, together with preliminary information on color models; Section 3 describes how the problem has been modeled in our research; Section 4 presents the experimental results of this study; in Section 5 conclusions are drawn and future directions outlined.

2 Color palettes

Color palettes are arrays of colors. Several color models exist, each aimed at describing colors as tuples of numbers (typically three or four values), called color components. RGB and CMYK are well known color models. In RGB, a color is described by three primary components R=red, G=green, B=blue. A color is obtained by additively combining intensities of the primary components. Vice versa in CMYK, primary components are C=cyan, M=magenta, Y=yellow, B=black, and colors are obtained by a subtractive aggregation of components. Both RGB and CMYK models are good for describing how to produce or print colors through devices. But besides them, there are also other models describing how colors are perceived. For this purpose, the CIE (International Commission on Illumination) introduced the XYZ model in 1931. Despite its age, it is still widely used in practice, especially as a reference for converting colors from one model into another. Similarly to RGB, XYZ adopts a system of additive primary components, namely X, Y, and Z. Each of these components represents the power perceived when RGB primaries are emitted. However, XYZ does not represent the response of cones at short, middle and long wavelengths. To better address the human perception of colors, in 1976 the CIE introduced the the CIELab model. In this model, color components are $L*$ that is a measure of color luminance, $a*$ being its position of red/magenta and green, and $b*$ its position between yellow and blue. When $L* = 0$, color is equivalent to black, whilst $L* = 100$ describes white. Uniform variations of components in the CIELab model aim at corresponding to uniform variations in the perception of colors. Therefore, this color model is suitable for measuring the perceptual distance between colors by means of the Euclidian distance ΔE between points in $L* \times a* \times b*$, that is

$$\Delta E = \sqrt{(L*_1 - L*_2)^2 + (a*_1 - a*_2)^2 + (b*_1 - b*_2)^2} \tag{1}$$

The maximum distance ΔE^* is between green and blue values. Distance ΔE provides a measure of both hue and density changes. According to recent studies, average observers are able to notice differences above $\Delta E = 5$ or 6, whilst a trained eye can notice differences from $\Delta E = 3$ or 4. The human eye, however, is much more sensitive to changes in gray levels and mid-tones; in that case differences of 0.5 delta-E may be noticed. The advantage of the CIELab model resides also in it being device-independent, thus resulting in more objective measures of colors. The same combination of $L*$, $a*$ and $b*$ will always describe exactly the same color. Color spaces are generally homeomorphic. Thus, there are formulas for transforming a color representation in one model, into an equivalent representation in another model.

The other key metrics we consider is the color contrast between contiguous colors. The W3C's WCAG [13] defines the *contrast ratio* as

$$C = \frac{\max(L_1, L_2) + 0.05}{\min(L_1, L_2) + 0.05} \tag{2}$$

where L is the *relative luminance*. Relative luminance is defined as the relative perceived brightness of any point, normalized to 0 for black and 1 for maximum white. We notice that relative luminance L as defined by W3C's WCAG differs from luminance $L*$ defined in CIELab. Contrast ratios can range from 1 to 21 (commonly written 1:1 to 21:1). According to W3C to reach level AAA of accessibility:" *text (and images of text) must have a contrast ratio of at least 7:1, except if the text is pure decoration. Larger-scale text or images of text can have a contrast ratio of 5:1*". This means that solution utility passing the contrast ratio threshold T_c is $C_u = 1$, decreasing below T_c. Therefore we adopted:

- RGB, for describing the palette colors
- CIELab, for measuring the distance between colors
- CIEXYZ, as the means for transforming RGB into CIELab

3 Interactive Genetic Algorithm in User Interface Design

Designers usually use guidelines to organize the layout and the features of user interface. Recent developments use meta-heuristic and evolutionary techniques [4] to organize structural elements of interface [8, 9] or non structural features such as colors [6]. Using meta-heuristics in finding the right interface configuration leads to the following advantages:

- A larger number of alternatives can be explored, thus pro-actively supporting human creativity and decision-making
- Different quality attributes and guidelines can be considered one at a time, thus facilitating the trade-off between conflicting criteria
- Designers are free to focus on more value-adding tasks, leaving algorithms to fine-tune their choices
- Interfaces can be automatically adapted to a larger set of devices, and a more specific set of user preferences

Ichikawa et al.[5] describe the re-working of Web pages for color-deficient viewers. The authors investigated a genetic algorithm able to improve the image contrast as perceived by simulation of impaired users, but still preserving the image chromaticity. To reach their goal, they first decompose the page into a hierarchy of colored regions. These spatial relations determine important pairs of colors to be modified. Then they minimize the fitness function using a genetic algorithm.

In a previous work [6] we confirmed Ichikawa et al. results proving the advantages in using a genetic algorithm for optimizing a color palette with a good trade-off between aesthetics and accessibility requirements.

In this work we introduce the user into the evolution process, as suggested by Takagi in Interactive Evolutionary Computation [10]. In particular we develop an IGA in order to include human preferences and knowledge into a Genetic Algorithm (GA). Indeed IGAs provide a mapping from users psychological space to GAs parameter space and thereby combine the power of human subjective evaluation with evolutionary computation [10].

In literature, recent works use IGAs for user interface design, in particular Quiroz et al. [8] encode user interfaces as individuals and run through a number of generations in order to support the exploration of user interface option space. Oliver et al. [9] focus their attention on the appearance and layout of Web sites. The user drives the evolution of style and layout of a Web pages through an IGA by picking solutions he/she prefers.

4 Algorithm

Users differ for age, education, skills, gender, and vision disorders. Therefore data cannot be used directly for evolving the interface, if we are interested to adapt the interface to specific needs and preferences.

An IGA application cannot rely on directly involving the user in large populations and hundreds of generations (like a canonical GA) because it is unrealistic to request a user to make hundreds or thousands of choices. If pressed for too much feedback, users are likely to lose interest and get tired [7].

According to the work developed by Quiroz [8], we do not invoke the user feedback at every generation on each individual, but at every t generations on some population representatives, as shown in Fig.1. This in order to minimize user interaction and to guarantee the user attention at the same time. The user feedback in collected just after the algorithm starts so to drive the initial evaluation of individuals during the subsequent iterations.

In other words, we aim at gaining the benefits of using a conspicuous number of individuals evolving along many generations, as in a canonical GA [3], but also to avoid an onerous user evaluation of a large number of palette. To reach this goal we used a clustering technique to group individuals and to identify their representatives to submit to the user judgment. The other individuals are evaluated according to their similarity to representatives. In particular we implemented the Expectation-Maximization (EM) algorithm as suggested in [12]. The EM algorithm was introduced by Dempster et al. [11].

In EM clustering, the algorithm iteratively refines an initial cluster model to fit the data, determining the probability each element can be associated to a cluster. The algorithm ends when the probabilistic model fits the data. The function used to determine the fit is the log-likelihood of the data given the model.

In details, EM is made of two steps until convergence:

- E-Step: estimate the Expected value of the unknown variables, given the current parameter estimation

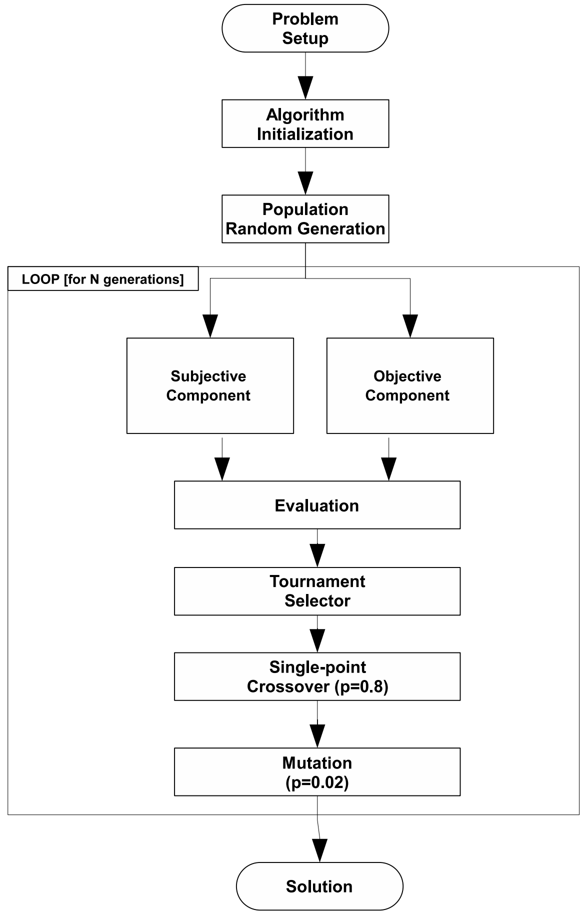

Fig. 1. The Interactive Genetic Algorithm

– M-Step: re-estimate the distribution parameters to Maximize the likelihood of the data, given the expected estimates of the unknown variables.

The result of the EM clustering method is probabilistic: unlike deterministic clustering techniques (e.g. K-means), a data point does not belong only to one cluster, instead it is associated to multiple clusters with different probability.

In Fig.1 we show the proposed algorithm, providing in Fig.2 and in Fig.3 the details of the two fitness components as expressed in Eq.5 and Eq.6.

The chosen clustering algorithm provides a means to partition the genetic population into a predefined number of clusters each of those being represented by an average individual, so the user is only asked to provide one feedback per cluster. Therefore, the user evaluates k palettes (where k is the fixed number of clusters). Individuals are affected by the cluster feedback proportionally to the probability of belonging to it.

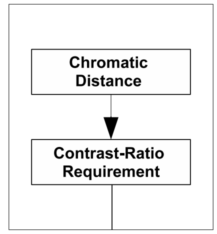

Fig. 2. Objective Component

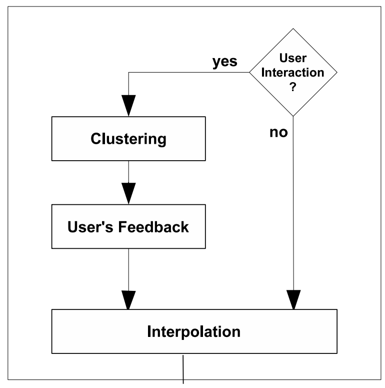

Fig. 3. Subjective Component

4.1 Model

The model for representing colors having been chosen, the palette chromosome coding is straightforward, as depicted in Fig.4.

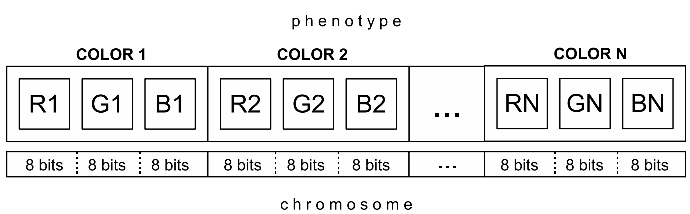

Fig. 4. Chromosome structure and phenotype mapping.

In particular, the chromosome is a bit string, reserving 24 bits (8 bits per component) to represent each color in the RGB space.

4.2 Fitness

Our aim is to find a palette that improves the luminance contrast between contiguous colors, while still preserving the original chromatic setting and awarding the palettes which best fit the user's preferences. This is done considering a fitness function made by an objective component depending on color distances and contrast ratios and a subjective component depending on the score given by user.

$$fitness = (objective^{w_1} \cdot subjective^{w_2})^{\frac{1}{w_1+w_2}} \tag{3}$$

where w_1 is the weight assigned to the objective component, whilst w_2 is the weight assigned to the subjective component weight; *objective* is the fitness objective component and *subjective* is the subjective fitness component. We used values of 1 for w_1 and w_2 for the experiments discussed in Section 5. Therefore, the fitness function, we adopted, is the following:

$$fitness = \left(\prod_{i=1}^{n} (1 - d_i) \cdot \prod_{j=1}^{m} c_j \cdot \prod_{h=1}^{n} (\sum_{g=1}^{k} p_g \cdot s_g) \right)^{\frac{1}{2 \cdot n+m}} \tag{4}$$

$$objective = \left(\prod_{i=1}^{n} (1 - d_i) \prod_{j=1}^{k} c_j \right)^{\frac{1}{2 \cdot n+m}} \tag{5}$$

$$subjective = \left(\prod_{h=1}^{n} (\sum_{g=1}^{k} p_g \cdot s_g) \right)^{\frac{1}{2 \cdot n+m}} \tag{6}$$

where d_i is the distance of resulting color i from the original one, c_j the contrast ratio of the m pairs of contiguous colors, p_g the probability that the particular palette belong to cluster g (with g = 0,1,...,k) and s_g the relative score. In particular,

$$d_i = \frac{\mathit{\Delta} E_i}{\mathit{\Delta} E^*} \tag{7}$$

$$c_j = \frac{20 + \min(C_j - T_j, 0)}{20} \tag{8}$$

where $\mathit{\Delta} E_i$ is the distance of color i from original one as defined in Eq.1, C_j is the contrast ratio as defined in Eq.2 and T_j is the contrast threshold (i.e. 7 or 5) as recommended by W3C's guidelines. It results that $f, d_i, c_j \in [0, 1]$. The maximum fitness value is $f = 1$, but this value is ideal as it is only reachable when $c_j = 1$, $d_i = 0$ for all i, j and this individual is evaluated by end-user as the best ($s_g = 1$), leading to the conclusion that the original palette is already optimal, thus not requiring any variation. In general, this value is below 1 because there is a need to vary colors ($d_i > 0$ for some i), some contrasts are below the threshold ($c_j < 1$ for some j), or the solution does not satisfy the user.

5 Experimental Results

In order to obtain a performance measure able to consider human in the loop, we make some simplifying assumptions and we define agents which simulates the human behavior by proving feedback according to some preference scheme. In particular, we defined two different agents:

– *AgentBY*, prefers palettes containing at least one color similar to blue, and one to yellow.
– *AgentCVD*, simulates the palette according to a specified color vision deficiency

AgentBY provides a score to the palette containing blue and yellow according to distance d as defined in Eq.9.

$$d = \frac{1}{2} \cdot (d_{blue} + d_{yellow}) \tag{9}$$

where d_{blue} and d_{yellow} are defined as:

$$\begin{aligned} d_{blue} &= \min\{d_{b_1}, d_{b_2}, ..., d_{b_c}\} \\ d_{yellow} &= \min\{d_{y_1}, d_{y_2}, ..., d_{y_c}\} \end{aligned} \tag{10}$$

and d_{b_i} and d_{y_1} are respectively the distance of color i from intended blue and yellow. Therefore AgentBY assigns a higher score to palettes having an average lower distance from yellow and blue.

AgentCVD assumes the user is color-vision impaired and simulates how he/she perceives a given color palette. In particular, we simulated a protanope user according to Viénot's model [13] and already tested in our previous work [6]. We compute the contrast ratio (Eq.2) between contiguous colors (e.g. foreground and background colors of a text area). Differently from the fitness colorimetric component, aimed at assessing the contrast of real colors, the agent provides a feedback according to the contrast of colors as perceived by an hypothetical protanope user.

In our experimentation, the agent feedback is collected on 5 palette color representatives, chosen as centroids of EM clustering.

Experimental results are presented in Fig.5. The algorithm use 1000 generations on populations made of 500 individuals. Charts refer to the average behavior of 10 different runs at different interaction rates. Fig.5(a) compares the performance of the fitness with different feedback rates ($F = 10$, $F = 50$,$F = 100$, and $F = 200$), simulated by AgentBY, while Fig.5(b) the performance for a different user behavior modeled by AgentCVD.

We can note there is a fitness improvement due to the explicit agent feedback. Improvement happens at iterations where the feedback is collected and it is stronger in the immediate following generations, highlighted by steps in the fitness graphics (Fig.5). Finally, the advantage of interacting with users is made clear by the overall value reached by the fitness function, that is lower when no feedback is collected during the algorithm evolution (about 25-30% lower than

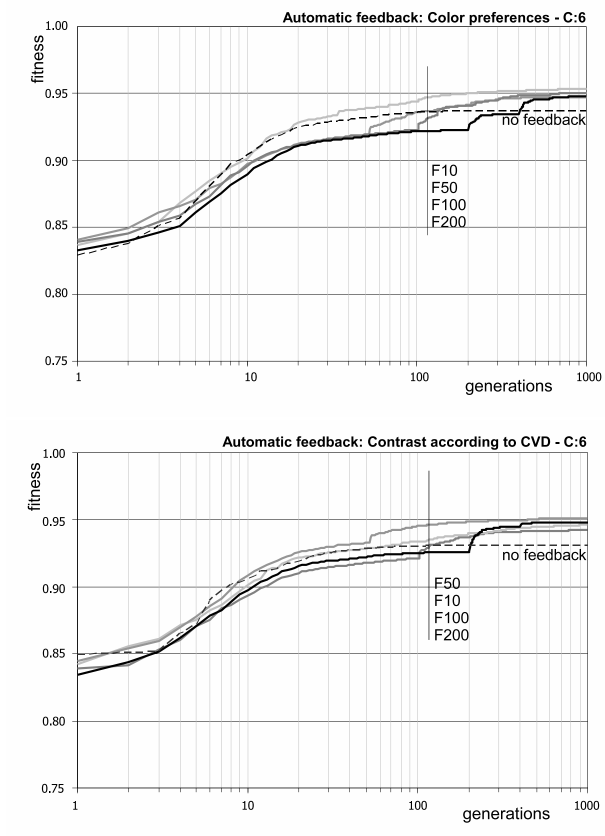

Fig. 5. Fitness provided by AgentBY (a) and AgentCVD (b)

the improvement given by the IGA from the initial fitness value). Indeed the algorithms reach a stable value after about 100 generations, this result is improved because of the user interactions. All these considerations emphasize the role of interactions in driving a genetic algorithms when part of the solution fitness is not explicitly modeled by the algorithm and it depends on user's preferences.

Our experimentation outlines that the overall result is independent on how frequent are user interactions, as all IGA configurations were able to reach a similar result, although more frequent feedbacks generally led to a little bit better score. As we assume that feedbacks are expensive to collect, due to the user fatigue, this should suggest that sparse interactions are desirable to more frequent ones.

Experimental results are confirmed for both agents, suggesting a general behavior of IGAs in this class of problems, as depicted in Fig.6.

As an example of application we can consider an initial palette of N colors c_n and the proposed algorithm provides a final accessible palette as a solution.

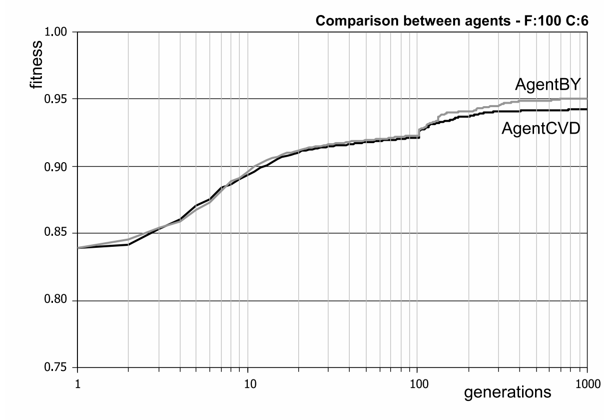

Fig. 6. Comparison of two different agents which simulate a specific human behavior every 100 generations.

Starting with a set of 6 colors whose representation is expressed in Eq.11 we optimize the colors according a particular model of correlation among colors. The adopted model correlates the first color (c_1) with the second (c_2) and the third (c_3), and the fourth (c_4) with the fifth (c_5) and the sixth (c_6).

The algorithm was setup with standard parameters according to Table 1.

Tournaments	1
Crossover probability	0.8
Mutation probability	0.02
Elitism	5

Table 1. Algorithm's parameters.

During the evolution process, the user evaluates 5 palettes every 100 generations and assigns to it a score between 1 and 5.

In Eq.12 we present the result and we observe how (c_1, c_2, c_3), whose contrast ratio is $2.3 : 1$ and $1.3 : 1$ with c_3, evolves into (g_1, g_2, g_3) with contrast ratio of $7.0 : 1$ and $5.8 : 1$. Furthermore c_4, which has a contrast ratio of 1:1 with c_5 and 2.7:1 with c_6 evolves into g_4, g_5 and g_6 with a contrast ratio of $5.7 : 1$ and $10.8 : 1$ respectively.

Furthermore the obtained solution satisfies human requirements about look and feels, in fact, in solution shown in Fig.7, we can observe how the hue of color c_6 is changed in g_6.

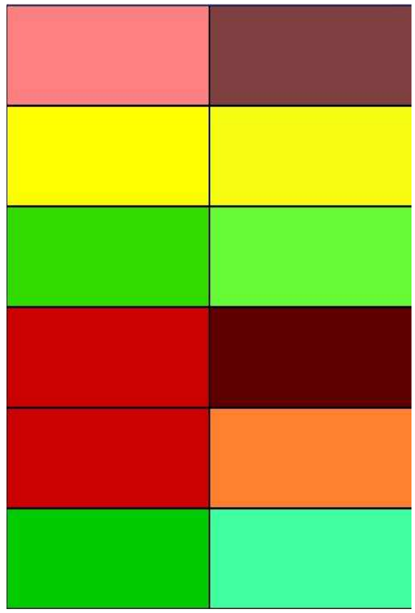

Fig. 7. Optimization of the original palette by means of IGA with human user feedback.

$$\begin{aligned} c_1 &= (255, 128, 128) \\ c_2 &= (255, 255, 0) \\ c_3 &= (51, 220, 0) \\ c_4 &= (204, 0, 0) \\ c_5 &= (204, 0, 0) \\ c_6 &= (0, 204, 0) \end{aligned} \tag{11}$$

where color c_i (with $i = 1...N$) is represented by the RGB components.

$$\begin{aligned} g_1 &= (144, 42, 47) \\ g_2 &= (255, 255, 0) \\ g_3 &= (128, 255, 59) \\ g_4 &= (63, 0, 0) \\ g_5 &= (255, 111, 50) \\ g_6 &= (14, 208, 13) \end{aligned} \tag{12}$$

Although the initial palette is unreadable for most of the users, thanks to the proposed algorithm the palette not only satisfies accessibility requirements, but also finds a good trade-off among W3C's requirements, chromatic choices, as planned by the interface designer, and user's preferences. We show in Eq.13 the values of the fitness for the optimization problem.

$$fitness_{manual\_feedback} = 0.959898 \tag{13}$$

6 Conclusions and Future Work

We presented an IGA that combines both computable metrics, taken from style guidelines, and human subjective feedback to drive the choice of an appropriate set of colors in a Graphical User Interface. Experimental results proved this approach to be feasible and advantageous, with respect to non-interactive genetic algorithms. This is due to the IGA ability of capturing fitness attributes related to human perception that cannot be acquired by mathematical modeling. In order to make analysis robust and independent on the human perception, we used agents to depict an expected behavior. Analysis showed that all IGA were able to reach a similar final result, independently on the interactions rate. This suggests that lesser frequent interactions are desirable on more frequent ones, as each interaction is expensive because of the user fatigue in evaluating a large number of solutions. In this paper we grouped individuals by EM clustering, and asking the user to evaluate a small number of solutions obtained as cluster centroids. Future directions are aimed at investigating if this is a robust approach, or there exist other methods able to group solutions in a way closer to the human perception of colors.

References

1. D. R. Commission. *The Web: Access and Inclusion for Disabled People.* April 2004. ISBN 0-11-703287-5.
2. J. D. M. Franoise Viénot, Hans Brettel. Digital video colourmaps for digital video colourmaps for displays by dichromats. *Color Research and Application*, 24(4):243–251, 1999.
3. D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison Wesley, 1989.
4. G. Russo, C. Birtolo, and L. Troiano Generative UI design in SAPI project In *CHI 08: CHI 08 extended abstracts on Human factors in computing systems*, New York, NY, USA, 2008. ACM.
5. M. Ichikawa, K. Tanaka, S. Kondo, K. Hiroshima, K. Ichikawa, S. Tanabe, and K. Fukami. Web-page color modification for barrier-free color vision with genetic algorithm. In *GECCO*, volume 2724 of *Lecture Notes in Computer Science*, pages 2134–2146. Springer, 2003.
6. L. Troiano, C. Birtolo, and M. Miranda. Adapting palettes to color vision deficiencies by genetic algorithm. In *GECCO '08: Proceedings of the 10th annual conference on Genetic and evolutionary computation*, pages 1065–1072, New York, NY, USA, 2008. ACM.
7. X. Llorá , K. Sastry, D. E. Goldberg, A. Gupta, and L. Lakshmi. Combating user fatigue in igas: partial ordering, support vector machines, and synthetic fitness. In *GECCO 05: Proceedings of the 2005 conference on Genetic and evolutionary computation*, pages 1363–1370, New York, NY, USA, 2005. ACM.

8. J. C. Quiroz, S. J. Louis, A. Shankar, and S. M. Dascalu. Interactive genetic algorithms for user interface design In *IEEE Congress on Evolutionary Computation, 2007. CEC 2007.* , pages 1366–1373. September 2007, IEEE.
9. A. Oliver, N. Monmarché, and G. Venturini. Interactive design of web sites with a genetic algorithm. In *Proceedings of the IADIS International Conference WWW/Internet*, pages 355–362, Lisbon, Portugal, november 13–15 2002.
10. H. Takagi. Interactive evolutionary computation: fusion of the capabilities of ec optimization and human evaluation. *Proceedings of the IEEE*, 89(9):1275–1296, Sep 2001.
11. A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.
12. T. Minka. Expectation-maximization as lower bound maximization, 1998. Tutorial published on the web at http://www-white.media.mit.edu/ tpminka /papers/em.html.
13. Web content accessibility guidelines 2.0. Technical Report 11, W3C, December 2007.
14. S. Zuffi, C. Brambilla, G. Beretta, and P. Scala. Human computer interaction: Legibility and contrast. In R. Cucchiara, editor, *ICIAP*, pages 241–246. IEEE Computer Society, 2007.

Acknowledgments

This work was partially supported by MIUR Project Automatic System For The Visually Impaired (SAPI), n.1642-2006.