MULTI-DIMENSIONAL OPTIMIZATION OF VISUAL CONTENT

Rosemarie J. E. Rajae-Joordens¹ and Ingrid Heynderickx¹,²

1. Philips Research Laboratories, Department of Visual Experiences, High Tech Campus 34, 5656 AE Eindhoven, The Netherlands
2. Delft Technical University, Faculty Electronic Engineering, Mathematics and Informatics, Mekelweg 4, 2628 CD Delft, The Netherlands

ABSTRACT

In order to give consumers the ability to adapt a display to their own image preference, current displays are equipped with a number of knobs, by which the viewer can tune various image quality attributes to his or her preference. Despite of this possibility, many people seem to avoid adjustments, because they do not know how to end up with a satisfying result. Results of two studies on preferred television settings, which we conducted in our laboratory in order to improve our products, clearly illustrate the difficulties many people encounter when optimizing image preference in a multi-dimensional parameter space spanned by various display settings.

The aim of the first study was to define a number of user profiles that might serve as predefined settings in future TV sets. Therefore, we investigated how a large group of users adjusted brightness, contrast, sharpness, saturation and white point of a conventional LCD-TV. In total, 176 participants started from a very poor image quality and were asked to adjust the TV to their own preference. Most people were unable to do so, and ended up with an image that was not pleasing them. In a second part of the study, the task was simplified: another 200 participants started from an acceptable image quality and had to further refine it by means of a series of one-dimensional optimizations. In this way, substantially more people were able to optimize the image to a satisfying result.

In the second study, participants were asked to adjust gain and offset of a Philips nine-view lenticular auto-stereoscopic 3D display to their preference. The gain setting affected the total range of stereoscopic depth that could be displayed, whereas the offset setting affected what part of the depth range was displayed in front or behind the screen. At first, 24 participants were requested to tune both gain and offset simultaneously. This resulted for many participants in inconsistent results. Therefore, the task was simplified. Another 21 participants were asked to tune the gain for some fixed values of the offset, and the offset for some fixed values of the gain. This largely reduced the spread in the data.

The results of both studies clearly show that in general people are not able to optimize display settings when navigating in a multi-dimensional space. It seems that the simultaneous optimisation of even as few as two dimensions is already too difficult. This may be a consequence of (1) the limited ability of humans to acquire a correct mental map of the multi-dimensional space, or of (2) limitations in the visual memory, though which viewers are unable to accurately recall the quality of the previous image in the adjustment process. As a result, participants tend towards the middle of a scale, even if this position on the scale is not an acceptable setting.
INTRODUCTION

In many contexts, image quality is erroneously considered to be equivalent to image fidelity, i.e. the ability to reproduce a real-world scene on an imaging medium, such as a photographic print or a display, as realistic as possible (Silverstein et al., 1996). Image fidelity, of course, is of high importance in contexts of e.g. security surveillance and medical imaging. For entertainment applications, such as television and home theatre, however, image quality should be more related to image preference, i.e. the degree of excellence of an image to the opinion of a viewer (Engeldrum, 2004). Image preference intrinsically is individual: what is preferred by one viewer is not necessarily also preferred by another viewer. For example, people from the Eastern-Pacific region seem to prefer more chromatic images as compared to Americans, and Japanese people show a tendency towards lighter images as compared to Europeans and Americans (Fernandez et al., 2002). Such differences are not only found over countries, but certainly also between individuals within a given population (Fernandez et al., 2002). To cope with this issue, current displays offer consumers the possibility to personalize the quality of the rendered image by adjusting some display settings according to their own preference.

Figure 1: Part of the user manual of the PHILIPS X26K176 colour television produced in 1972.
Already in earlier days, analogue televisions were equipped with a number of knobs, by which the viewer could tune features, such as brightness, contrast and sharpness (see Figure 1). Over time, more and more features were added and the consumer ended up with a number of menus and a remote control offering a large variety of options to adjust the quality of audio and video. Being able to tune the screen resolution, refresh rate, colour depth and white-point are just a few of the available options to alter the perceived image quality to a personal preference. Nowadays, in the digital era, the possibilities to adjust image quality is even further extended by offering the consumer the possibility to edit some of the signals going to the display, e.g. by means of photo editing software.

After purchase, however, many people just take the newly bought display out of the box and start using it without personalizing its performance by adjusting the quality related settings (Brouwer-Janse et al., 1992). What they often do not realize is that the pre-programmed factory settings are optimized for a shop environment, characterized by a high ambient illumination and a short viewing distance. Once used in the living room, characterized by a significantly lower ambient illumination and a considerably longer viewing distance, however, these settings are no longer optimal. Despite of the possibility to change the settings towards their own preferred optimum, many people seem to avoid making any adjustment.

The reason why consumers do not adjust the settings of their display simply may be laziness or ignorance about the possibility to personalize the display settings. This hypothesis is supported by the observation (unpublished internal studies) that even in case consumers are offered the possibility to optimize the display settings by means of a number of pre-programmed profiles, they still hardly switch between these profiles, but instead just select one, and stick to that. Based on some of our experimental results, however, we found a second reason why consumers in general do not adjust their TV settings: most of them are unable to find their own optimum in a parameter space containing two or more variables that can be changed simultaneously. To explore this hypothesis further, the results of two formerly performed experimental studies are studied again.

The aim of the first study was to define on a conventional LCD-TV a number of user profiles based on the value of five display settings, i.e. brightness, contrast, colour saturation, white point and sharpness. The resulting user profiles would be used as a basis for a wizard-like setup procedure in new television sets. In the second study, the optimal gain and offset setting in a Philips nine-view lenticular auto-stereoscopic 3D display were investigated. The gain setting affected the total range of stereoscopic depth that could be displayed, whereas the offset setting affected what part of the depth range was displayed in front or behind the screen. Both studies started from a protocol, in which participants were asked to adjust all the relevant parameters simultaneously, which implies that they had to navigate through a five-dimensional space in the first study and in a two-dimensional space in the second study. During both experiments, it became clear that the task was too difficult. Many participants in the first study indicated not to be satisfied with their end result, while in the second study a very large spread in data was found. As a consequence, the protocols were modified for both studies:
instead of a multi-dimensional optimization, the task was simplified into a series of one-dimensional optimizations at the time. In this way, substantially more people were able to optimize the image to a satisfying result in the first study, while the spread in the data was largely reduced in the second study. Thus, both studies supported our hypothesis that many people encounter difficulties when having to adjust more than one display setting at the time. At the end of this chapter, we will try to find explanations for this phenomenon.
STUDY 1: USER PROFILES

Various marketing studies have shown that image quality is one, if not the most important aspect for consumers to decide which television set to buy. As a consequence, optimizing image quality of television sets is of great importance to manufacturers. What complicates matters is that image quality is to a large extent a personal aspect (Fernandez et al., 2002). Hence, optimization of image quality can only be achieved by taking personalisation of TV settings into account. Some set makers support their customers in the process of personalizing the quality of their television by providing pre-programmed user profiles. Via one button on the remote control, consumers can select such a user profile, consisting of a unique combination of display settings. In this way, consumers are able to find TV settings that approach their own preference without too much effort.

In order to evaluate whether the proposed user profiles are properly selected, and indeed cover the optimal TV settings for the whole consumer population, the preferred television settings of a large group of people has to be collected. To this end, an experiment is performed in the Philips employee shop in Eindhoven, The Netherlands. Participants are recruited from the shopping audience, and are asked to adjust the contrast, brightness, colour, sharpness and tint setting of a LCD-TV set to their own preference.

As the preferred user profile is expected to depend on some participant characteristics, simple demographic values as age and gender are taken as co-variants in the experimental design. Also the favourite TV channel and favourite TV program genre, as well as the age of the TV at home are integrated into the experimental design as co-variants. The age of the TV at home is assumed to reflect the interest of the participant for new technology, although we realize that other factors, such as economic status, also may play a role in purchase behaviour.

Apart from the effect of individual on preferred user profile, also the dependency of image material is examined. The variation in video material used in the experiment mimic the daily life situation, including both broadcast and DVD recorded video material. As the broadcast video is intrinsically noisier than the DVD recorded material, and since the content of the videos is different, especially in terms of the ambient environment in which the depicted scene is captured (dark environment versus bright environment), an effect of image material on preferred user profile is expected.

MATERIAL AND METHODS

Participants
In total, 376 visitors of the Philips employee Shop in Eindhoven, The Netherlands, participated in this study. The employee shop was open to all
Philips employees and their relatives. The participants (297 males and 79 females) varied in age between 14 and 83 years. Note that there were no constraints with respect to vision and colour blindness.

**Experimental Setting**

A Philips 76 cm LCD-Wide screen Pixel Plus TV (type 30PF9975) with a LCD WXGA Active Matrix TFT screen (1280 x 768 pixels) was used in this study. The TV was placed on a 51.5 cm high TV table in a quiet corner in the back of the shop, and was connected to a standard DVD player via an S-Video cable. On both sides of the display, a halogen floor lamp was placed in order to create a domestic atmosphere and to obtain an ambient illumination of 25 Lux as measured on the centre of the screen in the direction of the viewer. A comfortable design chair was placed in front of the display at a distance of 2.40 meters (see Figure 2).

![Figure 2: Impression of the experimental setting.](image)

In this particular television set, brightness, contrast, colour saturation, white point and sharpness could be adjusted via the menu items brightness, contrast, colour, tint and sharpness, respectively. It should be noted that the brightness setting mainly adapted the black level of the television, while the contrast setting mainly determined the difference between that black level and the white level. The predefined user profiles implemented in this particular television were: “Rich”, “Natural”, “Soft”, “Multimedia” and “Eco”. As mentioned earlier, they each were
related to a unique combination of display settings, more particularly for the brightness, contrast, colour and sharpness setting. The values for the display settings programmed for a given user profile are summarized in Table 1. The user profile “Rich” corresponded to the most pronounced settings with a maximal value for contrast, and a high value for colour and sharpness. The user profile “Natural” is considered as a reasonable average with intermediate settings for contrast, colour and sharpness. The only difference between the user profile “Natural” and “Soft” is the different sharpness setting. The user profile “Multimedia” was a combination of a relatively low contrast, colour and sharpness value. The user profile “Eco”, the legally mandatory energy saving mode of the display, resulted in the least pronounced, but still acceptable image with a rather low value for contrast and colour and an intermediate value for sharpness. For all the user profiles, the brightness setting (mainly determining the black level of the television) is set at the same value.

<table>
<thead>
<tr>
<th>User profile</th>
<th>Contrast [0-100]</th>
<th>Brightness [0-100]</th>
<th>Colour [0-100]</th>
<th>Sharpness [0-7]</th>
<th>Tint [Cool, Warm, Normal]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rich</td>
<td>100</td>
<td>17</td>
<td>87</td>
<td>4</td>
<td>Not fixed</td>
</tr>
<tr>
<td>Natural</td>
<td>89</td>
<td>17</td>
<td>80</td>
<td>3</td>
<td>Not fixed</td>
</tr>
<tr>
<td>Soft</td>
<td>89</td>
<td>17</td>
<td>80</td>
<td>2</td>
<td>Not fixed</td>
</tr>
<tr>
<td>Multimedia</td>
<td>78</td>
<td>17</td>
<td>77</td>
<td>2</td>
<td>Not fixed</td>
</tr>
<tr>
<td>Eco</td>
<td>67</td>
<td>17</td>
<td>66</td>
<td>3</td>
<td>Not fixed</td>
</tr>
</tbody>
</table>

**Image material**

Two scenes, named “Harry Potter” (DVD quality, 50 sec, 50i, 16:9) and “Soap Series” (broadcast quality, 36 sec, 50i, 4:3), were prepared in a way dependent on the source. The MPEG2 encoded image material retrieved from DVD was simply captured without any further processing. The broadcast material, on the other hand, was first captured using an A/D converter (AJA video D10 AD), and subsequently encoded at a bit-rate of 8 Mbit/s with a MPEG2 encoder (TMPGEnc Plus 2.5). In order to facilitate presentation of the image material, a DVD per scene containing a continuous replay loop of that scene was burned. As a result, the DVD had to be restarted by the test leader only once an hour. In order to get an impression of the image material, one frame of each scene is shown in Figure 3.
Procedure

Initial Protocol

Prior to the experiment, a pilot experiment was performed in which experienced colleagues had to optimize the five TV settings simultaneously. As they did this in a fast and satisfying way, we decided to use the protocol in the experiment. Visitors were approached in the shop and asked if they were willing to participate in a short experiment. If so, they were requested to take place on the comfortable chair in front of the LCD-TV. The settings of the LCD-TV were adjusted extremely badly on purpose (contrast = 25; brightness = 25; colour = 25; sharpness = 0; tint = normal; colour-enhancement = on; dynamic-contrast = medium; digital-options = Pixel+; picture-format = auto; DNR = medium; active-control = off). On the LCD screen one of the two scenes as mentioned in the previous section was displayed.

Before starting the test, the participants’ gender, age, favourite TV channel (11 options), favourite TV program genre (9 options), and the age of their current TV at home were noted on a test form. Then, they were requested to adjust the five basic display settings, namely brightness, contrast, colour, sharpness and tint of the LCD-TV according to their personal preference via the menu using the remote control. If the participant asked for help, the test leader suggested a process of eight steps: 1) adjustment of contrast, 2) adjustment of colour, 3) adjustment of tint, 4) fine-tuning of colour, 5) adjustment of brightness, 6) fine-tuning of contrast, 7) fine-tuning of brightness, and 8) adjustment of sharpness. The test was concluded by having a look at all predefined user profiles present in the television (i.e. “Rich”, “Natural”, “Soft”, “Multimedia” or “Eco”) and indicating which one they liked the most. Once the participant had left the experimental setting, his or her personalized contrast, brightness, colour, sharpness and tint settings were filed together with the name of their favourite user profile. In total, 176 participants were tested using the initial protocol; 83 of them saw “Harry Potter”, the other 93 saw “Soap Series”.

Figure 3: Snapshots of the scenes “Harry Potter” and “Soap Series”.
Modified protocol

During the experiment, it became clear that non-experienced consumers were not satisfied with their final personalized optimization of the TV settings. The majority of them indicated to prefer the quality related to their favourite user profile above the quality resulting from their personalized TV settings. Therefore, we decided to change the experimental procedure by simplifying the task. First, participants had to indicate which of the five predefined user profiles they preferred. Subsequently, they had to optimise the TV further; this time not starting from very bad initial settings, but from the settings corresponding to their preferred user profile. Further optimization was forced along 5 steps only in a fixed order, each optimizing one dimension at the time, i.e. 1) tint, 2) colour, 3) contrast, 4) brightness, and finally 5) sharpness. Finally, participants were asked to compare the end result, i.e. the quality of their personalized TV settings with the quality corresponding to their favourite user profile, and to indicate which one they preferred. In total, 200 participants were tested according to this modified protocol; half of them saw “Harry Potter”, and the other half saw “Soap Series”.

Data Analysis

The statistical tests were performed using SPSS for Windows, version 11.5 (SPSS Inc., USA). Level of significance was taken to be 5%. First, overall mean brightness, contrast, colour, sharpness and tint setting were analysed using a multivariate analysis of variance (MANOVA) with Protocol (two levels: Initial and Modified) as between-subject factor. Because the expected main effect of Protocol was found, additional univariate analyses of variance (ANOVAs) were conducted to unveil the effect of Protocol on each of the display settings. All subsequent analyses were performed on the Initial and Modified Protocol data set separately.

For the Initial Protocol data set, it was evaluated whether the inability of people to adjust the LCD-TV according to their own preference was dependent on demographic factors. To this end, the data set was analysed by means of a MANOVA on brightness, contrast, colour, sharpness and tint setting with Scene (two levels: “Soap Series” and “Harry Potter”) as between-subject factor and age, gender, favourite TV channel, favourite TV program genre, and age of the current TV as covariates. In case of a significant factor or covariate effect, additional univariate ANOVAs were performed.

The Modified Protocol data set was further analyzed in order to determine whether the display settings adjusted by participants that were satisfied differed from those that were not satisfied, taking the demographic factors and the judged scene into account. To do so, a MANOVA on brightness, contrast, colour, sharpness and tint setting with Satisfaction (two levels: “user profile better than, or equal to own settings” and “user profile worse than own settings”) and Scene (two levels: “Soap Series” and “Harry Potter”) as between-subject factors, and age, gender, favourite TV channel, favourite TV program genre, and age of the current TV as covariates was performed. In case of a significant factor or covariate effect, additional univariate ANOVAs were carried out.
Finally, in order to identify possible effects of Gender on Satisfaction in the Modified Protocol data set, the non-parametric Mann-Whitney U test on Gender (two levels: male and female) by Satisfaction (two levels: “user profile better than, or equal to own settings” and “user profile worse than own settings”) was performed.

RESULTS

A MANOVA on preferred brightness, contrast, colour, sharpness and tint with Protocol (two levels: Initial and Modified) as between-subject factor revealed that the settings resulting from the initial protocol were significantly different from those obtained with the modified protocol [F(5,370)=69.50; P<0.001]. This finding was as expected, because the participants performing the experiment following the initial protocol complained that they were not satisfied with their personalized TV settings, whereas two-third of the participants following the modified protocol were satisfied with their final achievement. This finding implicates that a large group of people is not able to adjust their TV to a satisfactory level when starting from an extremely bad initial setting and/or having to optimize this setting, while freely or guided navigating in a five-dimensional space.

Figure 4: Visualization of the two data sets obtained via the initial and modified protocol in a 3D scatter-plot with the three relevant display settings (preferred contrast, brightness and colour) on the three axes.
Additional ANOVA’s showed that Protocol significantly affected preferred contrast \[F(1,374)=247.56; \ P<0.001\], brightness \[F(1,374)=74.93; \ P<0.001\] and colour \[F(1,374)=74.01; \ P<0.001\], but not tint \[F(1,374)=3.41; \ p=0.066\] and sharpness \[F(1,374)=0.25; \ p=0.615\]. In Figure 4, the two data sets are visualized in a 3D scatter plot with the three relevant display settings (i.e. preferred contrast, brightness and colour) on the three axes. As can be seen, the cloud of black markers, representing the data obtained with the initial protocol, is substantially shifted as compared to the cloud of open markers, representing the data obtained with the modified protocol. Therefore, further analyses of the data were performed on the two data sets separately.

**Initial Protocol**

As already stated above, a large number of participants spontaneously indicated not to be satisfied with their personal adjustment of the TV settings as a result of having followed the initial protocol. In order to get an idea whether the inability of people to adjust the LCD-TV was dependent on personal characteristics such as age or gender, a MANOVA on brightness, contrast, colour, sharpness and tint setting with Scene (two levels: “Soap Series” and “Harry Potter”) as between-subject factor, and age, gender, favourite TV channel, favourite TV program genre, and age of the current TV as covariates was performed. No significant effects of gender \[F(5,165)=1.53; \ p=0.183\], age \[F(5,165)=0.78; \ p=0.564\], favourite TV program genre \[F(5,165)=2.00; \ p=0.081\], favourite TV channel \[F(5,165)=0.22; \ p=0.954\] and age of current TV \[F(5,165)=0.64; \ p=0.667\] were found. Only the effect of scene \[F(5,165)=3.14; \ p=0.010]\) was significant. Further analysis showed that colour \[F(1,169)=4.68; \ p=0.032\] and tint \[F(1.169)=8.34; \ p=0.004\], but not contrast \[F(1.69)=0.00; \ p=0.949\], brightness \[F(1.169)=0.02; \ p=0.902\] and sharpness \[F(1.169)=1.70; \ p=0.194\] were responsible for the significant effect of scene. These findings suggest that the colour and tint setting are dependent of image content. Because only two different scenes were tested, the results were not sufficient to draw general conclusions about the effect of image content on tint and colour settings. For that reason, no further attention was paid to the separate scenes in this section, and results were averaged over the two scenes instead. This brings us to the conclusion that the inability of people to adjust a TV when starting from an extremely bad initial setting and having to navigate freely in a multi-dimensional space is not related to age, gender, favourite TV program genre, favourite TV channel or age of the current TV of the participant, whereas image content might play a role.

**Personalized TV Settings**

Although one might doubt to what extent the output of the adjustment task reflects the preference of the participants, because many participants mentioned not to be satisfied with their personalized TV settings, it is still interesting to have a look at the mean setting and standard deviation (S.D.) for contrast, brightness, colour and sharpness (see Table 2). Because no mean could be calculated for the nominal variable tint, percentage of preference was computed for each
possible item, resulting in “Cool” (1.7 %), “Normal” (45.5 %) and “Warm” (52.8 %).

Table 2: Overview of the minimum, maximum and mean brightness, contrast, colour and sharpness setting and standard deviation (S.D.) obtained via the initial protocol.

<table>
<thead>
<tr>
<th>Display Setting</th>
<th>Range</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contrast</td>
<td>[0-100]</td>
<td>32</td>
<td>100</td>
<td>61</td>
<td>15</td>
</tr>
<tr>
<td>Brightness</td>
<td>[0-100]</td>
<td>11</td>
<td>90</td>
<td>45</td>
<td>15</td>
</tr>
<tr>
<td>Colour</td>
<td>[0-100]</td>
<td>39</td>
<td>100</td>
<td>67</td>
<td>13</td>
</tr>
<tr>
<td>Sharpness</td>
<td>[0-7]</td>
<td>0</td>
<td>5</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

User Profiles

In Table 3, the preference expressed in numbers and percentages for the five user profiles is summarized. As can be seen, “Natural” and “Soft” were the two most preferred profiles. A closer look reveals that these two profiles had an identical contrast, brightness and colour setting; they only marginally differed in the sharpness setting. Ignoring this difference in sharpness setting for a moment, we can conclude that 65 % of the participants preferred a contrast setting of 89, a brightness setting of 17 and a colour setting of 80. Note, however, that these values were not necessarily the most optimal values for contrast, brightness and colour in terms of personal image quality, because the participants were limited in their choice by only five predefined sets of settings. Nonetheless, the results in Table 3 illustrate which values for the settings best approached the participants’ personal optimum.

It has to be noted that the naming of the user profiles might have influenced the results. Several participants mentioned to associate the name “Multimedia” with a PC- or game application, and so, not to be appropriate for a television setting. Similarly, the name “Natural” might have had a positive cognitive association with image quality and “Eco” with an ecologically sound way of living. It is not clear to what extent the naming of the user profiles has affected the results, but definitely, its impact should not be neglected.

Table 3: Preference scores expressed in quantity (N) and percentage for the five user profiles.

<table>
<thead>
<tr>
<th>User Profile</th>
<th>Characteristics</th>
<th>Display Settings</th>
<th>Preference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Contrast [0-100]</td>
<td>Brightness [0-100]</td>
<td>Colour [0-100]</td>
</tr>
<tr>
<td>Rich</td>
<td>100</td>
<td>17</td>
<td>87</td>
</tr>
</tbody>
</table>
Personalized TV Settings versus user profiles

Figure 5 shows the averaged contrast, brightness and colour setting as adjusted by the participants following the initial protocol compared to those of the two most favourite user profiles “Natural” and “Soft”. It demonstrates that the personal TV settings (given by the black bars) tend toward values near the middle of the scale; that is near a value of 50. The value of the contrast, brightness and colour setting for the most favourite user profiles (given by the grey bars) are located more towards the extremities of the scale. Thus, it seems that the participants tended to adjust the TV settings towards the middle of the scale.

<table>
<thead>
<tr>
<th>Display Attribute</th>
<th>CONTRAST</th>
<th>BRIGHTNESS</th>
<th>COLOUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Setting (+SD)</td>
<td>80</td>
<td>40</td>
<td>66</td>
</tr>
</tbody>
</table>

**Figure 5:** Overview of the mean contrast, brightness and colour setting (+ standard deviation) adjusted by the participants (black bars) and the contrast, brightness and colour of the most favourite user profiles, i.e. “Natural” and “Soft” (grey bars).

Modified Protocol

In the last step of the modified protocol, participants compared the quality related to their favourite user profile with the quality resulting from their personalized TV settings, and indicated which one they preferred. This result was recorded as Satisfaction, and was given two values, i.e. “User profile better than, or equal to
own settings” and “User profile worse than own settings”. The first step was to determine whether the personalized TV settings were significantly different for those persons that judged their result to be satisfactory compared to those that judged their result not satisfactory, taking the demographic data and the judged scene into account. In case no such difference would be found, the factor satisfaction could be further ignored. Therefore, a MANOVA on brightness, contrast, colour, sharpness and tint setting with Satisfaction (two levels: “User profile better than, or equal to own settings” and “User profile worse than own settings”) and Scene (two levels: “Soap Series” and “Harry Potter”) as between-subject factors, and age, gender, favourite TV channel, favourite TV program genre, and age of the current TV at home as covariates was performed. No significant effect of gender \[F(5,187)=1.81; p=0.114\], favourite TV program genre \[F(5,187)=0.26; p=0.936\], favourite TV channel \[F(5,187)=1.06; p=0.387\], age of current TV \[F(5,187)=0.48; p=0.790\], scene \[F(5,187)=2.05; p=0.074\]) and satisfaction \[F(5,187)=0.75; p=0.585\] was found. Also, the interaction between scene and satisfaction \[F(5.187)=1.17; p=0.325\) was not significant. Only the covariant age showed a significant effect \[F(5,187)=4.72; P<0.001\]. Thus, only age affected the way people adjusted the LCD-TV.

Additional ANOVAs showed that age significantly affected the personalized TV settings of brightness \[F(1,191)= 4.97; p=0.027\], colour \[F(1,191)=6.32; p=0.013\] and sharpness \[F(1,191)=12.02; p=0.001\], but not of tint \[F(1,191)=0.09; p=0.093\] and contrast \[F(1,191)=0.14; p=0.714\]. After removing only 3 outliers with a brightness setting of more than 65, however, the significant effect of the covariate age on brightness completely disappeared \[F(1,188)=1.87; p=0.173\]. The significant effect of the covariate age on colour vanished \[F(1,182)=2.74; p=0.100\] after excluding the 9 participants preferring the tint “Cool” from the data set. In Figure 6, a scatter plot shows the relationship between age on the X-axis and the adjusted sharpness setting on the Y-axis. The positive slope of the regression line indicates that as age increased, a higher sharpness setting was preferred. In general, the process of aging is accompanied by a decreased visual acuity, so it is not surprising that higher-aged persons tend to compensate the loss in detail and contour vision with a higher sharpness setting. However, it should also be noted that many participants reported difficulties in distinguishing the perceived sharpness associated with different sharpness settings, and indicated that, therefore, they just selected one of the medium values. Because the effect of the covariate age on the preferred brightness, colour and sharpness setting was anyway weak and not fully understood, it was ignored in the remainder of the analysis.
Figure 6: Relation between age and preferred sharpness setting, and the regression line defined by the equation $\text{Preferred Sharpness Setting} = 2.70 + \text{Age} \times 0.03$.

Personalized TV Settings

In Table 4, a summary of the mean ($\pm$ S.D.) personalized optimum for the contrast, brightness, colour and sharpness setting is shown. In addition, the percentage of preference was computed for each possible tint value, resulting in “Cool” (4.5 %), “Normal” (50.5 %) and “Warm” (45.0 %). A comparison of these results with those obtained via the initial protocol (see Table 2) shows that the mean setting for contrast and colour was considerably higher when using the modified protocol than when using the initial protocol. The mean setting for brightness obtained with the modified protocol was considerably lower than that obtained with the initial protocol.

<table>
<thead>
<tr>
<th>Display Setting</th>
<th>Range</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contrast</td>
<td>[0-100]</td>
<td>58</td>
<td>100</td>
<td>81</td>
<td>10</td>
</tr>
<tr>
<td>Brightness</td>
<td>[0-100]</td>
<td>3</td>
<td>80</td>
<td>33</td>
<td>14</td>
</tr>
<tr>
<td>Colour</td>
<td>[0-100]</td>
<td>54</td>
<td>99</td>
<td>77</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 4: Overview of the minimum, maximum and mean preferred brightness, contrast, colour and sharpness setting and standard deviation (S.D.) obtained via the modified protocol.
Table 5 summarizes the preference scores for the five predefined user profiles. Identical to the results obtained with the initial protocol, the user profiles “Natural” and “Soft” were the two most preferred ones. It should, however, be noted that the difference in preference between the second (i.e. “Soft”) and third (i.e. “Rich”) most preferred user profile was less than 2% as obtained with the modified protocol. This was substantially smaller than what was found with the initial protocol; there the difference in preference between the user profile “Soft” (i.e. second most preferred) and “Rich” (i.e. third most preferred) was as large as 10.8%. Nonetheless, it can be concluded again that 65 % of the participants preferred a contrast setting of 89, a brightness setting of 17 and a colour setting of 80 in case of a forced choice out of five predefined sets of settings.

Table 5: Preference scores expressed in quantity (N) and percentage for the five predefined user profiles.

<table>
<thead>
<tr>
<th>Characteristics Display Settings</th>
<th>Preference</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Profile Contrast [0-100] Brightness [0-100] Colour [0-100] Sharpness [0-7] Tint [Cool, Warm, Normal]</td>
<td>N %</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Rich 100 17 87 4 Not fixed</td>
<td>38 19.0</td>
</tr>
<tr>
<td>Natural 89 17 80 3 Not fixed</td>
<td>89 44.5</td>
</tr>
<tr>
<td>Soft 89 17 80 2 Not fixed</td>
<td>41 20.5</td>
</tr>
<tr>
<td>Multimedia 78 17 77 2 Not fixed</td>
<td>23 11.5</td>
</tr>
<tr>
<td>Eco 67 17 66 3 Not fixed</td>
<td>9 4.5</td>
</tr>
</tbody>
</table>

Personalized TV Settings versus user profiles

In Figure 7, the averaged contrast, brightness and colour setting as adjusted by the participants following the modified protocol (see black bars) are compared to the settings of the two most favourite user profiles (see grey bars). As can be seen, the mean contrast and colour setting adjusted by the participants is close to the value predefined for the settings of the “Natural” and “Soft” user profile. The latter is not true for the mean brightness setting, which shows a substantial tendency towards the middle of the scale for the value adjusted by the participants.

Asking the participants at the end of the experiment to compare their personalized TV settings with their preferred user profile taught us that the majority (54.5 %) was satisfied with their personal settings, 39 % preferred their favourite user profile, and the other 6.5 % saw no difference. More than half of the females (54 %) were dissatisfied with their personalized TV settings against a third of the males (36 %) only. An explanation might be that females lacked self-
confidence with regard to their performance on such a (technical) assignment. A large number of female participants at first refused to participate in the test. They argued that their husband was responsible for the adjustment of the TV at home, and that they had no clue how to do it. A Mann-Whitney U test on Gender by Satisfaction revealed a significant gender effect (p=0.042). In other words, females more negatively judged their personalized TV settings as compared to males despite the fact that they had adjusted the LCD-TV in a similar way.

![Figure 7: Overview of the mean contrast, brightness and colour setting (+S.D.) adjusted by the participants (black bars) and the contrast, brightness and colour of the most favourite user profiles, i.e. "Natural" and "Soft" (grey bars).](image)

SUMMARY

The aim of the present study was to define a number of user profiles on the basis of preferred TV settings. Therefore, an experiment was conducted in the Philips employee shop in which a large number of people were asked to adjust a LCD-TV set according to their preference.

After having tested 176 participants, i.e. about halfway the planning, it became clear that many people were not able to adjust the TV according to their own preference when starting from an extremely bad initial setting. This inability was not related to the used image content or age, gender, favourite TV program genre or favourite TV channel, or age of the current TV of the participant. Because participants got lost in the five-dimensional space (contrast, brightness, colour, sharpness, and tint), they tended to select a value at the middle of the scale, even if this was far from an optimal setting.
Therefore, a modified, and at the same time simplified experimental protocol was defined for another 200 participants. Instead of starting from a very bad initial setting, they started from their most preferred predefined user profile, which they had to optimise further in 5 subsequent one-dimensional steps. This time, the majority of participants were satisfied with the end result.

Remarkably, however, satisfied and dissatisfied participants did not differently adjust the LCD-TV. Moreover, they preferred the same user profile. Relatively more females claimed to be dissatisfied with their own adjustments. These results suggest that dissatisfied participants lacked self-confidence with regard to their performance on such a (technical) assignment.

It has to be kept in mind that the difference in contrast and colour setting, and to a lesser extent the difference in brightness setting between the most preferred user profile and the averaged personal adjustment obtained via the modified protocol was rather marginal. Since comparing the results of the personal adjustment with the favourite predefined user profile could only be done temporally by sequentially scanning through the complete list of user profiles, the marginal differences might have been very difficult to judge due to limitations of our visual memory system. It cannot be excluded that if participants would have had the possibility to compare the image resulting from the personal adjustment of the settings to the image corresponding to the preferred user profile side-by-side, a more pronounced or potentially different preference might have resulted.

The participant characteristics age, gender, favourite TV program genre, favourite TV channel and interest in new technology did not influence how the LCD-TV was adjusted. Therefore, a set of user profiles that were generally applicable were defined and integrated in new Philips TV sets.
STUDY 2: GAIN AND OFFSET

Philips has developed a multi-view auto-stereoscopic 3D display. The term auto-stereoscopic refers to the fact that stereoscopic depth (i.e. a different perspective of the scene for the left and right eye) can be observed without goggles. This is achieved by means of a transparent sheet of lenses mounted in front of a standard LCD (Van Berkel and Clarke, 1997; Willemsen et al. 2006; Redert et al. 2006). The lenses are designed such that sets of pixels are projected in different neighbouring viewing directions, hereafter referred to as views. This implies that each view might contain a slightly different image; in practice, this is a slightly different perspective of a given scene. The basic concept is illustrated in Figure 8, and the numbers in this figure refer to the seven views generated by the lenticular sheet. As a consequence, at any horizontal position at a given viewing distance in front of such a lenticular 3D display, the viewer sees a different image with his left eye and right eye. These two images are combined in the brain into a stereoscopic 3D experience (Coren and Ward, 1989, p.291-294), which means that the viewer might be able to perceive objects hanging in front or behind the display screen. Moreover, the viewer can move in front of the screen and still experience stereoscopic depth, and even motion parallax, which is the possibility to “look around” objects. Since the cone existing of the seven different views is repeated in different directions in the horizontal plane, stereoscopic depth can be perceived by multiple viewers at the same time.

The multi-view auto-stereoscopic 3D display, commercialized by Philips, has nine views. It is found to be strongly preferred over a 2D display by a majority of video gamers, because it provokes stronger feelings of presence during gaming (Rajae-Joordens et al., 2005; Rajae-Joordens, 2008). Feelings of presence are investigated using psychophysiological measures, i.e. galvanic skin response and heart rate (to measure arousal) in combination with subjective measures, i.e. questionnaires (to measure the valence of the experience). Also a trend towards an increase in feelings of presence and emotions is found in people watching TV on a 3D display (Rajae-Joordens, 2008), and again a considerable number of participants prefer the 3D TV above the 2D TV (Seuntiëns et al, 2007; Kaptein et al. 2008). The difference in strength of the experience between the two applications can be partly explained by the difference in the source of the 3D content used: the gaming application uses perfectly graphically rendered 3D material, while for the TV application video material has to be converted from 2D to 3D due to the fact that 3D content is nowadays insufficiently available.
Figure 8: Principle behind a multi-view auto-stereoscopic lenticular 3D display as designed by Philips. By means of lenses mounted in front of a standard LCD slightly views are generated, i.e. seven views in this specific example. As a consequence, the viewer sees a different image with his left eye and right eye. These two images are combined in the brain into a stereoscopic 3D experience.

Conversion from 2D to 3D is done by estimating for every pixel in the 2D image a depth value based on cues in the image content (see e.g. Tam and Zhang, 2006). Such cues may be occlusions, luminance differences, sharpness gradients, etc. Conversion then results in a 2D image (containing per pixel the required luminance and chromaticity values) and a depth map (containing per pixel the estimated depth value). These two pixel maps are combined in a so-called 2D-plus-depth video format (Redert et al. 2007). Figure 9 shows an example of a 2D-plus-depth video format.

Figure 9: Example of the 2D-plus-depth video format, consisting of a 2D image (containing per pixel the required luminance and chromaticity values) and a depth map (containing per pixel the estimated depth value).
Subsequently, the 2D-plus-depth video signal is rendered on the multi-view lenticular 3D display by means of a real-time rendering algorithm. This rendering algorithm contains two important parameters affecting the depth perception: gain and offset. Gain defines the amount of depth in the scene. It is a measure for the amount of difference in the images of neighbouring views. Offset determines the positioning of the depth in front or behind the display screen. Both parameters are illustrated in Figure 10.

![Figure 10: Illustration of the parameters gain and offset in the rendering algorithm of 3D content.](image)

In order to optimize the 3D viewing experience, gain should be chosen such that the difference in the images of neighboring views is large enough to perceive stereoscopic depth, though small enough to avoid the visibility of annoying depth artifacts, such as blurring and ghosting. These artifacts may be a consequence of leakage of image content of one view into another, a phenomenon called “crosstalk” (Kooi and Toet, 2004; Seuntiëns, et al. 2005; Kaptein and Heynderickx, 2007), or of the inability of our brain to fuse two images with a two large difference in perspective (Yeh and Silverstein, 1990). The offset, on the other hand, should be chosen such that part of the image content, which is expected to be exciting, is suspended in front of the display screen. However, incomplete objects located on the edge of the screen might disturb the experience of increased realism, because they are only partially visible in front of the screen, and thus not natural or realistic (a phenomenon called “window violation”). Almost certainly, there will not be just one single optimum for the settings of gain and offset, but instead a display type and content type dependent optimum. Hence, it is the aim of this study to investigate the optimal values for gain and offset for different situations.

Optimal gain and offset values are expected to be dependent on the design of the lenses of the lenticular screen. For example, the distance between the lenticular screen and the display and the pixel pitch determine the width of the viewing cone. The Philips nine-view Extreme 3D display is designed such that it allows to reproduce a larger depth range, but at the expense of a narrower viewing cone. This implies that moving left and right in front of the display is more critical for artefacts. The Philips nine-view Comfort 3D display, on the other hand, has a broader viewing cone, and hence, allows more freedom in moving left and
right in front of the display without seeing artefacts, but has intrinsically a smaller depth range. In this study, the preferred gain and offset value is evaluated for these two types of 3D displays.

Video content itself is also expected to affect the optimal gain and offset value. Content captured with a 3D camera set-up or computer generated 3D content has more accurate depth information than 2D-to-3D converted content, and hence, usually contains fewer artefacts. Therefore, it is hypothesized that a higher gain value will be preferred for captured than for converted 3D content. Additionally, since changes in depth over time in displayed 3D video content require adaptations of the human eye in terms of accommodation and convergence, and hence may cause visual discomfort for large changes in depth at high speed (Yano et al., 2002; Emoto et al., 2004; Emoto et al., 2005), the preferred gain and offset value is expected to depend on the speed with which information in depth is varied over time. The presence of subtitles in video content may affect the optimal gain and offset value too, because in the air floating subtitles are perceived to be unnatural. Finally, based on the assumption that a black frame around the displayed video content may enhance depth perception, an effect of such a black frame on the preferred gain and offset value is expected as well. Based on all these expectations, different types of video content are used to investigate the optimal gain and offset value on two types of lenticular 3D displays.

**MATERIAL AND METHODS**

**Participants**

In total, 45 Philips employees were invited to participate in this study. The participants (36 males and 9 females) varied in age between 24 and 45 years.

**Experimental Setting**

Two Philips auto-stereoscopic nine-view lenticular 3D displays, i.e. a 37” Extreme and a 42” Comfort, both with a display resolution of 1920 x 1080 pixels and an optimal viewing distance of 3 meter were used in this study. The 37” Extreme 3D display had a field of view of 15.0°, a viewing cone of 6.2° and a horizontal pixel size of 0.4275 mm. The 42” Comfort 3D display had a field of view of 17.2°, a viewing cone of 21° and a horizontal pixel size of 0.4845 mm. The two displays were placed on a 80 cm high table, and with their backside close to a white wall. By means of adhesive tape, the floor was marked at a distance of 3 meter from of the displays to indicate the optimal viewing distance to the participants. The test room was illuminated at a level of 3 lux measured on the centre of the screens in the direction of the viewer.

Both displays were addressed with a PC, containing a GeForce video card. The Hydra board in the 3D displays allowed real-time rendering of the 2D-plus-depth input signal for a given combination of gain and offset values. The input signal had
a spatial resolution of 1920 x 540 pixels, and consisted of two side-by-side frames of 960 x 540 pixels each: the left one containing the MPEG compressed 2D image content and the right one containing the corresponding depth map (see Figure 8 for an example). Video input was rescaled to 1920x1080. By means of a wireless game pad (Xgear, NVidea XFX), participants were able to adjust the gain (range [0,63]) and offset (range [0,224]) settings of the real-time rendering algorithm. A gain of 0 corresponded to an image in 2D, whereas for a gain of 63 the full available depth range was used. For an offset of 0, all content was displayed in front of the screen; for an offset of 224, all content was displayed behind the screen. An offset of 122 implied that 50% of the depth range was displayed in front of and 50% behind the screen. In essence, the rendering algorithm calculated per pixel the appropriate screen disparity (i.e. the distance in pixels between corresponding content in two neighbouring views) from the depth value in the input signal, the gain and the offset value. Due to hardware limitations (related to the specific design of the lenticular screen) the maximal screen disparity that could be used, was limited. As a consequence, also the maximal depth range that could be rendered was restricted. Content that by the combination of input depth value, gain and offset should have been displayed beyond this limit was clipped to the border of the depth range.

Image material
Initial protocol
A computer generated 3D advertisement provided by Philips 3D Solutions (1200 frames, 50Hz) was used. This advertisement contained a logo that moved in and out the screen, while rotating. A snapshot of the content is given in Figure 11. In order to evaluate the effect of the presentation speed on preferred gain and offset, besides the original scene (“Logo Normal Presentation Speed”), two more scenes were generated by means of frame dropping: one consisting of 600 frames (“Logo Double Presentation Speed”) and another one of 300 frames (“Logo Quadruple Presentation Speed”).

![Figure 11: Snapshot of the scene "Logo".](image)
Modified protocol

Halfway the experiment, the protocol was changed due to the fact that simultaneously adjusting gain and offset turned out to be rather difficult for most participants. Because we already collected enough data on 3D generated image material and presentation speed, we decided to focus on 2D-to-3D converted content instead.

Five scenes, named “Spy”, “Spy Framed”, “Spy Subtitles”, “Spy 3D” and “Horse”, were prepared. The scene “Spy 3D” originated from 3D material in anaglyph format with the content for the right eye being red filtered and the content of the left eye being green filtered. It was converted to the 2D-plus-depth input format by filtering the images of both eyes. The other four scenes all originated from 2D material and were converted to the 2D-plus-depth input format with a real-time 2D-to-3D conversion algorithm based on a combination of spatial gradient cues, gravity, luminance and motion cues. The scene “Spy” contained an object that moved towards the viewer while fully remaining inside the screen, while the most important object in “Horse” did not completely remain inside the screen, but instead was located on the edge and thus was only partly visible. The scenes “Spy Subtitles” and “Spy Framed” were prepared out of the scene “Spy” by adding subtitles and black bars, respectively. The subtitles were placed in a front layer. Due to the characteristics of the conversion algorithm, the upper black bar was assigned to a rear layer in the image content, while the lower black bar was assigned to a front layer. Snapshots of the scenes used are shown in Figure 12.
Figure 12: Snapshots of the scenes “Spy” (a), “Spy Framed” (b), ‘Spy Subtitles” (c), “Spy 3D” (d), and “Horse” (e).

Procedure

Initial protocol

After entering the test room, the participants were asked to stand behind the 3 meter line on the floor in front of one of the two 3D displays. They were allowed to move freely as long as they stayed behind this line. They then received the game pad, on which the four shoulder buttons, placed along the edges of the pad, were programmed such that gain could be adjusted with the left hand (button 5 = increase gain value; button 7 = decrease gain value) and offset with the right hand (button 6 = increase offset value; button 8 = decrease offset value). Before starting the test, they got some time to investigate what happened when pushing the buttons. Subsequently, they were asked to adjust gain and offset such that they experienced the best performance of the display. They had to do this twice for each of the three scenes (“Logo Normal Presentation Speed”, “Logo Double Presentation Speed” and “Logo Quadruple Presentation Speed”): i.e. starting from an initial setting once with a minimal gain of 0 and an offset of 112, and once with a maximal gain of 63 and an offset of 112. After that, they were requested to repeat the procedure on the other display. Half of the participants started on the 37” Extreme 3D display; the other half of the participants started on the 42” Comfort 3D display. Note that participants could take as much time as they needed to find their optimal gain and offset settings. They did not receive any feedback on where on the scale of gain and offset values they were during the adjustment. In total, 24 participants performed this test, each in an individual session.

Modified protocol

During the experiment, it became clear that the task of simultaneously adjusting gain and offset was rather difficult for most participants. Hence, we simplified the task in the second part of this study by splitting it into two separate tasks. Another 21 participants were requested to tune the offset for two fixed gain values (i.e. 15 and 22) and to tune the gain for three fixed offset values (i.e. 87, 112 and 137) for each of the five selected scenes (“Spy”, “Spy Framed”, “Spy Subtitles”, “Spy 3D” and “Horse”). Because, however, the software tool to change gain and offset was limited to 24 stimuli only, we omitted the tuning of gain with an offset of 137 for the “Spy 3D” scene. Stimuli were presented in a random
order, and per tuning the initial value of the parameter to be adjusted alternately changed between its lowest and highest value (0 or 63 for gain, and 0 or 224 for offset). Due to logistic issues, most participants first performed the experiment on the 37” Extreme 3D display, followed by a second session on the 42” Comfort 3D display, roughly one week later.

Data Analysis

The statistical analyses were performed using SPSS for Windows, version 12 (SPSS Inc, USA). The level of significance was taken to be 5%.

The Initial Protocol data set was analyzed by means of two ANOVAs, i.e. on gain and offset separately with Presentation Speed (three levels: “Logo Normal Presentation Speed”, “Logo Double Presentation Speed” and “Logo Quadruple Presentation Speed”), Display (two levels: 37” Extreme 3D display and 42” Comfort 3D display), Order (two levels: First Extreme then Comfort 3D display and First Comfort then Extreme 3D display) and Initial Gain Setting (two levels: 0 and 63) as within-subject factors and Participant included as random factor. To investigate the effect of presentation speed, display and initial gain setting on preferred gain and offset. Apart from the interactions with the random factor Participant, all other 2-way interactions were included in the analysis. In case a post-hoc test was required, the Tukey's HSD Post Hoc Test was used.

Based on a boxplot representation, all outlying data points were first individually removed from the Modified Protocol data set. In practice, this implied that 54 adjustments out of the 1008 in total were not taken into account in the analysis. On the resulting Modified Protocol data set two ANOVAs were performed. In order to investigate the effect of content, display and offset setting on preferred gain, an ANOVA on gain with Scene (five levels: “Spy”, “Spy Framed”, “Spy Subtitles”, “Spy 3D” and “Horse”), Display (two levels: 37” Extreme 3D display and 42” Comfort 3D display) and Fixed Offset Setting (three levels: 87, 112 and 137) as within-subject factors and Participant included as random factor was executed. The second ANOVA on offset with Scene (five levels: “Spy”, “Spy Framed”, “Spy Subtitles”, “Spy 3D” and “Horse”), Display (two levels: 37” Extreme 3D display and 42” Comfort 3D display) and Fixed Gain Setting (two levels: 15 and 22) as within-subject factors and Participant included as random factor was done to examine the effect of content, display and gain setting on preferred offset. Apart from the interactions with the random factor Participant, all other 2-way interactions were included in the analysis. Again, in case a post-hoc test was required, the Tukey's HSD Post Hoc Test was used.

RESULTS

Initial protocol

Figure 13 shows 286 collected data points, each reflecting a combination of preferred gain and offset as indicated by the 24 participants twice (i.e. for the two initial settings) for the three scenes on the two displays. Due to a technical issue,
2 data points got lost. As can be seen in Figure 13, the data points were located all over the parameter space without any clear shape in the cloud of data points, suggesting that preferred gain and offset were largely spread and not related. The lack of a relation between offset and gain was further confirmed by a very low Pearson correlation coefficient ($r=0.071$; $p=0.229$). A closer look at the data revealed two distinct sources of spread: (1) a “between-subject” spread, and (2) a “within-subject” spread. The latter was caused by a number of participants that exhibited a wide spread in the preferred gain and offset combination over their twelve adjustments. This might for example have been because their preferred gain and offset was very sensitive to variations in presentation speed and display, or because they were not able to find their optimal setting in the two-dimensional parameter space. The first source of spread resulted from a large difference in preferred gain and offset combination among participants that each individually were rather consistent over their twelve adjustments. In order to control for between-subject differences and to deal with the absence of a relation between the two parameters, gain and offset settings were analyzed separately with Participant included as random factor.

![Figure 13: All preferred gain and offset combinations as adjusted by 24 participants.](image-url)
Gain

An ANOVA on gain with Presentation Speed (three levels: “Logo Normal Presentation Speed”, “Logo Double Presentation Speed” and “Logo Quadruple Presentation Speed”), Display (two levels: 37” Extreme 3D display and 42” Comfort 3D display), Order (two levels: First Extreme then Comfort 3D display and First Comfort then Extreme 3D display) and Initial Gain Setting (two levels: 0 and 63) as within-subject factors and Participant included as a random factor was performed. Significant effects of Display \[ F(1,249)=6.20; p=0.013 \], Presentation Speed \[ F(2,249)=3.19; p=0.043 \] and Participant \[ F(22,249)=14.5; p<0.001 \] were found, whereas neither a significant effect for Initial Gain Setting \[ F(1,249)=0.50; p=0.481 \], nor Order \[ F(1,249)=0.02; p=0.904 \] was found. Moreover, none of the two-way interactions was statistically significant.

In Figure 14, the mean preferred gain (+ SEM) is depicted for each presentation speed and for each display. It shows that a slightly smaller preferred gain (27.0 ± 1.5) was found at the 42” Comfort 3D display as compared to the 37” Extreme 3D display (30.8 ± 1.4).

In addition, the preferred gain for the normal presentation speed (average 26.8 ± 1.7) and quadruple presentation speed (average 31.1 ± 1.8) significantly differed \( p=0.040 \). No difference in preferred gain between the double (average 28.9 ± 1.9) and normal presentation speed \( p=0.468 \), as well as between the double and quadruple presentation speed \( p=0.416 \) were found. An explanation might be that due to faster speeds artefacts became less visible and as a consequence higher gain were preferred. From these results, it can be concluded that with increasing presentation speed the preferred gain slightly increased.
Offset
An ANOVA on offset with Presentation Speed (three levels: “Logo Normal Presentation Speed”, “Logo Double Presentation Speed” and “Logo Quadruple Presentation Speed”), Display (two levels: 37” Extreme 3D display and 42” Comfort 3D display), Order (two levels: First Extreme then Comfort 3D display and First Comfort then Extreme 3D display) and Initial Gain Setting (two levels: 0 and 63) as within-subject factors and Participant included as a random factor was performed. Offset was not found to be affected by Display \([F(1,249)=0.68; \ p=0.411]\), Presentation Speed \([F(2,249)=0.01; \ p=0.988]\), Order \([F(1,249)=0.34; \ p=0.559]\) and Initial Gain Setting \([F(1,249)=0.08; \ p=0.783]\). On the other hand, a significant effect of Participant \([F(22,249)=5.43; \ p<0.001]\) as well as a significant interaction effect between Display and Initial Gain Setting \([F(1,249)=9.62; \ p=0.002]\) and between Presentation Speed and Order \([F(2,249)=5.38; \ p=0.005]\) were found.

Because no main effect of presentation speed, display, order and initial gain setting was found, it is sufficient to mention that the overall mean preferred offset was 114.1 (± 2.9). This finding implied that participants preferred to have roughly
half of the depth range displayed in front of the screen and half behind the screen.

The significant interaction between Display and Initial Gain Setting resulted from the fact that starting at an initial gain setting of 63 instead of 0 yielded an increase in preferred offset of 14.1 for the 37” Extreme 3D display and in a decrease in preferred offset of 17.2 for the 42” Comfort 3D display (see Figure 15). The significant interaction between Presentation Speed and Order showed that starting the test on the 42” Comfort 3D display instead of on the 37” Extreme 3D display resulted in a reduction in preferred offset of 14.8 for the double presentation speed and in an increase in preferred offset of 24.9 for the quadruple presentation speed (see Figure 16). Both interaction effects indicated that the way the two displays were adjusted was influenced by the experimental design. But, as the between-subject differences were large due to the fact that gain and offset had to be adjusted simultaneously, the occurrence of these interaction effects should be interpreted with care.

![Figure 15: Mean preferred offset settings (+SEM) for two different initial gain settings on two different 3D displays.](image-url)
Modified protocol

Gain
An ANOVA on preferred gain with Scene (five levels: “Spy”, “Spy Framed”, “Spy Subtitles”, “Spy 3D” and “Horse”), Display (two levels: 37” Extreme 3D display and 42” Comfort 3D display) and Fixed Offset Setting (three levels: 87, 112 and 137) as within-subject factors and Participant included as random factor was executed. Significant effects of Scene \([F(4,511)=13.6; \ p<0.001]\), Display \([F(1,511)=10.5; \ p=0.001]\), Fixed Offset Setting \([F(2,511)=8.40; \ p<0.001]\) and Participant \([F(20,511)=28.4; \ p<0.001]\) were found, as well as a significant interaction effect between Fixed Offset Setting and Scene \([F(7,511)=4.38; \ p<0.001]\).

In Figure 17, mean preferred gain (+SEM) is depicted for each scene and for each display. Although hardly visible, a slightly larger gain was found for the 42” Comfort 3D display \((13.7 \pm 0.7)\) as compared to the 37” Extreme 3D display \((12.6 \pm 0.6)\).

A closer look at the effect of scene revealed that the preferred gain decreased in descending line from “Horse” (average 17.1 ± 1.1), over “Spy Framed” (average 13.5 ± 1.0), “Spy” (average 13.1 ± 0.9), and “Spy Subtitles” (average 10.8 ± 0.8), to “Spy 3D” (average 9.8 ± 0.9). The scene “Horse” was adjusted to a significantly higher preferred gain than all other scenes (all \(p<0.001\)). Within the Spy-based content, a significantly lower gain was found for the anaglyph-to-3D
converted “Spy 3D” scene as compared to the 2D-to-3D converted scenes “Spy” (p=0.013) and “Spy Framed” (p=0.003), but not as compared to “Spy Subtitles” (p=0.842). Participants indicated that they chose a lower gain setting for the “Spy 3D” scene because of annoying artefacts that occurred at higher gain values. Apparently, these artefacts were less visible in the 2D-to-3D converted scenes. Further, the preferred gain for the “Spy” scene did not significantly differ from the preferred gain for the “Spy Framed” (p=0.990) and “Spy Subtitles” (p=0.109) scenes. The latter two were adjusted to a significantly different preferred gain (p=0.033). In other words, adding subtitles or a frame to the content of a scene had no effect on the preferred gain.

Figure 17: Mean preferred gain settings (+SEM) on two different 3D displays and for the scenes “Horse”, “Spy Framed”, “Spy”, “Spy Subtitles” and “Spy 3D” (* p<0.01; # p<0.01 as compared to each other scene; + p<0.05 as compared to “Spy 3D”; & p<0.05 as compared to “Spy Framed”).

Figure 18 is another representation of the same data, now including the effect of fixed offset. It shows a significantly lower gain (11.4 ± 0.7) for an offset of 87 as compared to an offset of 112 (14.5 ± 0.8; p<0.001) or an offset of 137 (13.6 ± 0.9; p=0.008). The latter two offset values did not differently affect the preferred gain.
In other words, a significantly smaller gain was preferred when the depth range was shifted towards the front of the screen.

The significant interaction between fixed offset setting and scene on preferred gain was mainly caused by the difference in behaviour between the “Horse” scene and the “Spy” scene; for the “Horse” scene the preferred gain increased with offset value, whereas for the “Spy” scene the preferred gain decreased with offset value. The main difference in content between those two scenes was that the most important object in the “Horse” scene was located on the edge of the screen, and therefore only partly visible, whereas the scene “Spy” contained an object that moved towards the viewer while fully remaining inside the screen. Taking this aspect of the content into account, it was not surprising that the “Spy” scene was adjusted to a high gain in combination with the lowest offset, because in this way the object moving towards the viewer was most convincingly displayed in front of the display. For the “Horse” scene, the high gain in combination with the highest offset helped to minimize the impression of an unreal situation by displaying the partial content behind the screen.

![Bar chart showing mean preferred gain settings (+SEM) overall and for five different scenes as obtained via three different fixed offset settings.](image)

*Figure 18: Mean preferred gain settings (+SEM) overall and for five different scenes as obtained via three different fixed offset settings (*p<0.01 as compared to Fixed Offset settings 112 and 137).*
Offset
An ANOVA on preferred offset with Scene (five levels: “Spy”, “Spy Framed”, “Spy Subtitles”, “Spy 3D” and “Horse”), Display (two levels: 37” Extreme 3D display and 42” Comfort 3D display) and Fixed Gain Setting (two levels: 15 and 22) as within-subject factors and Participant included as random factor was performed. Offset was found to be significantly affected by Scene \(F(4,366)=5.01; p=0.001\) and Participant \(F(20,366)=2.67; p<0.001\), whereas no significant effect of Display \(F(1,366)=0.18; p=0.670\) and Fixed Gain Setting \(F(1,511)=1.45; p=0.229\), nor of any two-way interaction was found.

Because no main effect of display and fixed gain setting was found, it is sufficient to show the mean preferred offset for each scene (see Figure 19). As can be seen, the preferred offset decreased in descending line from “Spy 3D” \((122.6 \pm 6.1)\), over “Horse” \((116.9 \pm 7.3)\), “Spy Framed” \((104.2 \pm 4.1)\), and “Spy Subtitles” \((98.6 \pm 4.2)\) to “Spy” \((95.3 \pm 4.5)\).

In more detail, a significantly higher preferred offset was found for the “Spy 3D” scene as compared to the scenes “Spy” \(p=0.002\) and “Spy Subtitles” \(p=0.011\), but not as compared to the scenes “Horse” \(p=0.936\) and “Spy Framed” \(p=0.089\). Further, the scene “Horse” was adjusted to a significantly higher preferred offset than the scene “Spy” \(p=0.031\), but not than the scenes “Spy Framed” \(p=0.414\) and “Spy Subtitles” \(p=0.097\). From these findings, it could be concluded that the scenes “Horse” and “Spy 3D” were preferably positioned somewhat more behind the screen, while the scene “Spy” was preferably located somewhat more in front of the screen. This was as expected because the objects in the scene “Spy” were fully remaining in the screen, while the scenes “Horse” and “Spy 3D” contained partly visible objects located on the edge of the screen. Finally, the preferred offset for the “Spy” scene did not significantly differ from the preferred offset for the “Spy Framed” \(p=0.746\) and “Spy Subtitles” \(p=0.992\) scenes. The latter two were also not tuned to a significantly different preferred offset value \(p=0.941\). Thus, adding subtitles or a frame to the content of a scene had no effect on the preferred offset.
SUMMARY

The aim of the present study was to determine the optimal value for the gain and offset setting in two nine-view lenticular 3D Philips displays. Therefore, an experiment was conducted in which participants were asked to optimize gain and offset of a 37" Extreme 3D display and a 42" Comfort 3D display while viewing computer generated 3D content played at three different speeds.

After having tested 24 participants, it became clear that many people had problems with the simultaneous adjustment of gain and offset. Because a large spread in the data was found and the preferred gain and offset did not correlate, it could be concluded that many participants got lost in the two-dimensional parameter space. Therefore, we decided to define a modified, yet simplified experimental protocol. Another 21 participants were asked to tune the gain for some fixed values of the offset, and the offset for some fixed values of the gain instead of tuning the parameters simultaneously. In this new approach based on a series of one-dimensional tunings, the spread for gain (SEM became 0.4 instead of 1.0) and offset (SEM became 2.5 instead of 2.9) was reduced, and moreover, participants stopped complaining about the difficulty of the task, as...
they did in the two-dimensional tuning task of the initial procedure. Unfortunately, it was not possible to statistically compare the results of both protocols due to the use of different image material. Nonetheless, the data of the set of one-dimensional tunings were found to be more consistent, indicating that tuning in a two-dimensional parameter space is already rather difficult for participants.

Besides information on the ability of people to find an optimum in a one- or two-dimensional parameter space, the study also provided information on the actual preferred gain and offset. In summary, preferred offset was on average chosen such that roughly half of the content was displayed in front of the screen, and the other half behind the screen. This value was independent of the type of 3D display and the type of 3D source material (i.e. whether it was computer generated 3D or 3D converted content). The only variable that affected the preferred offset value was whether or not objects were partly covered by the borders of the display screen: the offset shifted with 11.3% from the screen towards the viewer for content, in which the main objects fully remained within the borders of the screen, whereas for content, in which the main objects were only partially visible, the offset shifted backwards with 6.9%. In addition, the averaged preferred gain was slightly different for the 37" Extreme 3D display than for the 42" Comfort 3D display, both when using computer generated 3D material or 2D-to-3D converted material. Further, as expected, preferred gain slightly increased with the presentation speed and decreased for image material, in which artefacts became annoyingly visible at a larger depth range. Finally, adding subtitles or a frame to the content of a scene had no effect on the preferred gain. These findings have been used to optimize an algorithm for automatic gain and offset adjustment on Philips 3D displays.

ABILITY TO OPTIMIZE IN A MULTI-DIMENSIONAL SPACE

From internal studies, we know that many consumers do not use the possibility to change the settings of their display towards their own preferred optimum. As mentioned in the introduction, laziness or ignorance of the possibility may be possible reasons for this behaviour. Based on more experimental evidence, however, another possible explanation becomes evident: most people are unable, or at least experience it as difficult, to find their own optimum in a parameter space containing two or more variables that can be changed simultaneously. To explore this hypothesis further, the results of two experimental studies are used.

The aim of the first study was to define on a conventional LCD-TV a number of user profiles based on the value of five display settings, i.e. brightness, contrast, colour, tint and sharpness. The purpose of the second study was to determine the optimal gain and offset for a Philips nine-view lenticular auto-stereoscopic 3D display. Both studies started from a protocol in which participants were asked to adjust all the relevant parameters simultaneously. This implied that they had to
navigate through a five-dimensional space in the first study, and through a two-dimensional space in the second study. During both experiments, it became clear that the task was too difficult. Many participants in the first study indicated not to be satisfied with their end result, while in the second study a large spread in the data was found. As a consequence, the protocol was modified for both studies: instead of a multi-dimensional optimization, the task was simplified into a series of one-dimensional optimizations. In this way, substantially more people were able to optimize the image to an according to themselves satisfying result in the first study, while the spread in the data was reduced in the second study. Thus, both studies supported our hypothesis that many people encounter difficulties when having to adjust more than one display setting at the time.

Simultaneous optimisation of even as few as two dimensions seems already to be too difficult. This may be a consequence of the limited ability of humans to acquire a correct mental map of a multi-dimensional space. A study on information seeking showed that individuals with a low spatial visualisation ability (i.e. reflecting a person’s ability to manipulate spatial patterns into other arrangements) took longer to complete a certain task and made more errors on first attempts to find information in hierarchical databases compared to those with a high spatial visualisation ability (Downing et al. 2005). Training and practice could improve this spatial visualisation ability (Vincent & Williges, 1988). The adjustment tasks used in our study required people to navigate through a complex multi-dimensional space instead of through a hierarchical system. The difference in task might have had effect on the way a mental representation of a space was formed, yet the effect of training and practice most probably remained comparable. Because most of the participants in our studies were not experienced in adjusting display settings, and additionally, got only limited opportunity to practice during the short period of the experiment, the inability to obtain a correct mental representation of a complex multi-dimensional space during a short period of time seems to be a reasonable explanation for the absence of a satisfactory and/or consistent result in adjusting display settings. This is further supported by the preliminary results of the pilot experiment in study 1: experienced colleagues were able to optimize the five TV settings simultaneously in a fast and satisfying way, probably by making use a correct mental map of the parameter space, obtained via long-term experience. Another spatial-navigation study demonstrated that acquiring spatial mental models from a map was relatively more time-consuming than from a route description (Brunyé and Taylor, 2008). Our participants, however, did not get any tool at all to facilitate the visualization of the multi-dimensional space; their only reference was the change in appearance and the name of the changed display setting. Thus, the above argumentation makes it reasonable to assume that the inability of people to adjust two or more display settings simultaneously is due to an inability to acquire in a short period of time a correct mental map of the multi-dimensional space necessary for an accurate navigation, especially in the case of inexperienced people.

Visual-spatial tasks, such as the mental rotation task, are typically better performed by men than women (Dabbs et al. 1998; Collins & Kimura, 1997). This
gender difference has been partially contributed to hormonal differences between males and females; males who produce little or are insensitive to androgens (e.g. testosterone), or have increased estrogen levels, show reduced spatial abilities, whereas females with high androgen (e.g. androstenedione) levels show greater spatial abilities (Coren and Ward, 1989, p. 522-524). Recently, it has been demonstrated that even pre-puberal boys show superior performance to similar-aged girls in the Virtual Morris Water task, in which participants have to navigate in a virtual reality environment in order to find a target in a virtual water pool, in analogy to the traditional tests often used in rodent spatial learning studies (Newhouse et al. 2007). This finding suggests that gender differences in learning how to navigate in a spatial environment exist prior to puberty, and do not require the effects of sex hormones at puberty. Instead, early-life hormonal effects or different preferential learning strategies by boys and girls may be responsible for the gender differences. Even at an early age, males and females may differ in the type of activities they engage in, as toys and games that have a spatial component, such as blocks, are typically more preferred by boys than by girls (Coren and Ward, 1989, p. 524). Finally, the results of a recent study suggest that the gender gap in mental rotation performance is partially caused by experiential factors, particularly those induced by socio-cultural stereotypes. Positive stereotype messages emphasizing a female advantage enhance the performance of females, while the same messages cause males to perform worse (Wraga et al. 2006). Thus, considering these findings, we would expect a gender difference in our studies too.

In contrast to this expectation, however, we found in study 1 that males and females adjusted the TV in a similar way, independent of whether they had to perform the five-dimensional task of the initial protocol or the series of one-dimensional tasks of the modified protocol. Unfortunately, a possible gender effect could not be reliably analysed in study 2 due to a shortage of female participants; because it was not part of the initial research question, no attention was paid to recruit an equal amount of males and females. Based on the results of study 1, the question arises why we did not find the expected gender difference. Two recent publications on gender studies may give an answer to this question. In the study already mentioned above (Newhouse et al, 2007), boys and girls initially needed an equal amount of time to find the target in the Virtual Morris Water experiment, but after already four trials boys became significantly faster than girls. A second study demonstrated that in a virtual hole-board test, in which participants had to learn a given path via positive feedback (again in analogy to a test developed to investigate spatial navigation in rodents), the error rate and travelled distance to solve the task was initially equal for males and females, but started to differentiate after only one trial (Cánovas et al. 2008). Thus, the performance in navigation initially seems to be the same for both sexes, but males seem to construct a correct mental representation faster and benefit from this more in subsequent trials than females. Taken this into account, it is not surprising that we did not find any gender effect on the way the TV was adjusted as participants adjusted the TV only once.
Despite the fact that males and females did not differently adjust the LCD-TV, and also preferred the same predefined user profiles, relatively more females than males claimed to be dissatisfied with their final adjustment result. As their performance was equal to that of males, it seems that the females simply lacked self-confidence with regard to their performance on such a technical assignment. Much research is done on gender differences in relation to technology. Most of this research focuses on computer use. It is demonstrated that females exhibit a tendency towards a more negative attitude to computers, lower computer self-confidence, a higher computer anxiety and a shorter use of the internet than males (Durndell et al. 2000; Durndell and Haag, 2002). A recent study, however, also shows that the gender gap is closing as far as computer access and self-confidence is concerned (Imhof et al. 2007). What should not be overlooked, however, is the fact that the participants of these studies were university students, while the participants in our study in the Philips employee shop varied considerably more in age and education. A large number of our male participants had a technical background, as they worked in a technical function within Philips, but many female participants seemed to suffer from technophobia as they refused to participate in our experiment at first. Their reason for refusing was that their husband was responsible for the technical tasks at home and that they had no clue how to adjust the TV settings. These remarks, often heard during the experiment, give evidence to a low self-esteem for these females with regard to the assignment. Already a long time ago, it was demonstrated that persons with a high self-esteem perceived themselves as doing better on a task than persons with a low self-esteem, even if their performance on a certain concept-formation task was comparable (Shrauger and Terbovic, 1976). This is exactly what happened in our study. Therefore, the feeling of dissatisfaction of females towards their own adjustment performance seems to be due to a lower self-confidence and technophobia.

So far, we discussed the inability to acquire a correct mental map of a multi-dimensional space as a possible explanation for the difficulties observed during the task of optimizing video content along multiple parameters simultaneously. However, memory may also have affected the inability to adjust more than one display setting at the same time. When changing a display setting, the image appearance changes too. At that moment, a decision has to be made whether the previous image is worse or better in quality than the subsequent one. The same problem occurs when comparing the favorite user profile to the personalized settings. Since comparing the results of the personal adjustment with the favourite predefined user profile could only be done temporally by sequentially scanning through the complete list of user profiles, the possibly marginal differences between both images may have been insufficient to reliably assess the most preferred one due to limitations of our visual memory system. It has for example been shown in literature that the human visual system is good at detecting small colour differences, and so, can differentiate millions of colours on condition that the colours can be compared side-by-side (Hunter et al., 1987); but, the number of distinguishable colours drastically decreases to only hundreds of colours for time-sequential colour matching, i.e. when the human vision
system has to rely on memory (Newhall et al., 1957; Uchikawa, 1983; Uchikawa et al., 1986; Pérez-Carpinell et al., 1998). Similarly, it has been demonstrated that the visibility threshold for a non-uniformity in hue across the display screen increases with a factor 3 to 4 when presenting the original and distorted image time-sequentially on the same display instead of simultaneously on two neighbouring displays (Heynderickx et al., 2002). The decrease in accuracy with which a new stimulus is matched to a sample, as a consequence of an increased time delay between stimulus and sample, is described in a more theoretical way by White et al. (White, 2002; White and Wixted, 1999). They describe a model that mainly attributes the decrease in accuracy over time to a process called diffusion, which reflects the increase in spread of locating the sample along its characteristic dimension (e.g. brightness, colour). When we requested the participants to compare the image resulting from their personalized settings to the image corresponding to their favourite predefined user profile, diffusion of the sample image (as described by White) may have played a considerable role, as the temporal distance between viewing the two images was relatively large, and distractive information in the form of non-relevant images, i.e. user profiles other than the preferred one, were shown intermediately. In the process of adjusting the TV settings, however, the effect of diffusion of the sample image most probably was less important as subsequent images were shown rapidly one after the other. Moreover, in case diffusion would have played a dominant role, we would not have seen a considerable improvement when changing the protocol from a simultaneous adjustment in a multi-dimensional space to a set of one-dimensional adjustments. In other words, memory may play a role in the inability of people to adjust two or more display settings simultaneously, but is certainly not the only factor responsible.

Interestingly, when people got “lost” in the multi-dimensional space, they tended at least in study 1, to adjust the TV settings towards the middle of the scale, even if this position on the scale did not yield an acceptable value. In study 2, participants had to adjust the gain and offset without any visual feedback of the scales, and therefore, a tendency towards the middle could not be found here. Instead, a scattering over the whole scale was seen. The tendency to use the middle of a scale regardless of content is a well known phenomenon, and is called “midpoint responding” in marketing research (Baumgartner and Steenkamp, 2006). It is found to be independent of cross-cultural or gender differences (Grimm and Church, 1999). The occurrence of midpoint responding strongly implies that the data obtained via the multi-dimensional adjustment of the TV settings should be treated with extreme care.

CONCLUSION

The results of two studies were used to investigate how people behave when requested to optimize video content along more than one parameter dimension simultaneously. In both studies the results for the simultaneous adjustment of the settings did not show much consistency, yielded a large spread, and the participants frequently reported to be dissatisfied with their final result.
Simplifying the task to a set of one-dimensional adjustments at the time reduced the spread in the data and improved the satisfaction of the participants with their result. The inability of people to adjust two or more display settings simultaneously may be due to the inability to acquire a correct mental map of the multi-dimensional space necessary for an accurate navigation. Especially in the case of inexperienced viewers that are only given a limited amount of time to get experienced, the lack of a correct representation of the multi-dimensional space may have a large impact on the ability to optimize the rendering of video content. Additionally, the limited performance of the human visual memory may play a role. The inability to find an optimum in a multi-dimensional space is in first instance not gender related, although females have the impression that they perform worse, which is probably due to a lower self-confidence and technophobia.
ACKNOWLEDGEMENTS

The authors want to thank Mark Lambooj, Dijana Bosnjak, Arnold van Keersop, Erno Langendijk, Franco Oberti, Stefan Swinkels, Elena Vicario and Ingrid Vogels for their dedicated role as test leader. We are also grateful to Marc Op de Beeck and his team for their help on 2D-to-3D conversion of video material and the Philips 3D Solutions team for their technical support.

References


