Abertay University

Modelling Image Quality for Automotive Display Technologies

By

Dorothee Christine Wolf, Dipl.-Ing. (FH), M.Sc.

A thesis submitted in partial fulfilment of the requirements of

Abertay University

For the degree of

Doctor of Philosophy

January 2014

I certify that this is a true and accurate version of the thesis approved by the examiners.

Signed…….. Date…….
(Director of Studies)

DECLARATION

Declaration of originality and Permission to copy

Author: Dorothee Christine Wolf

Title: Modelling Image Quality for Automotive Display Technologies

Qualification: PhD **Year of submission:** 2014

(i) The author certifies that the above-mentioned report is his own original piece of work.

(ii) The author gives the authorisation to consult and to copy parts of this work for personal use only. Any other use is limited by the Laws of Copyright. Permission to reproduce any material contained in this work should be obtained from the author.

Date: ……. **Signature:** ……..

ABSTRACT

The aim of this thesis is to link perceived image quality to physical display parameters. This is done in the context of automotive displays. Specialities of automotive display applications like high ambient lighting conditions and the necessity to access information quickly are explained. A summary of readability models relevant to automotive applications is given and the difference between readability and perceived image quality is explained. The methodology chosen to investigate perceived image quality is the Image-Quality-Circle framework by Engeldrum (2000). Engeldrum states that observers form their image quality rating by weighting the visual attributes they perceive. Visual algorithms, which can be investigated via psychometric scaling, link visual attributes to the underlying physical image parameters which are typically measure by physical instruments. The visual attributes investigated in this thesis are perceived contrast, brightness, blackness and colourfulness. Perceived contrast, brightness and blackness are derived from display luminance via the DICOM just noticeable difference (JND) scale. Colourfulness is scaled based colour gamut in the CIE1931 chromaticity diagram. It was shown that image quality rating rises with growing perceived contrast; the limiting factors are glare and perceived blackness. In colourfulness scaling a linear relationship between colour gamut and colourfulness rating was demonstrated. Higher colourfulness can compensate lower brightness in perceived image quality.

ACKNOWLEDGEMENTS

All experiments in this thesis were performed in the facilities of the BMW Group during a three year PhD placement within the display pre-development group. Special thanks go to Mr. Robert Isele and Mr. Michael Heimrath from BMW Group and to Prof. Karlheinz Blankenbach for making this thesis possible and for valuable support during my time at BMW Group.

The good angel of this thesis and person who finally coaxed it out is Dr. Colin Cartwright my principal supervisor. Dear Colin thank you for your patience and moral boosting over this long time. Not only did you fund additional experiments at the University Abertay Dundee which I performed during an unforgettable 4 month stay at the EPICentre in Dundee. Through all the ups and many, many long downs you were always accessible, always had good advice and never gave up on me. It was your belief that there is a PhD in me that really made this PhD finally happen.

Dear Christian without your support this PhD would not have happened either! You took over so many tasks for a really long time to give me enough time for writing up and always did your best to get me into a good thesis mood.

My parents are another part in not giving up altogether – I just did not want to explain to Daddy that he sold his model railway for nothing.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

The importance of electronic displays in cars has been constantly growing (Heimrath 2000). The first display to be introduced in a BMW was a seven segment digital clock in the year 1980. Since then more and more electronic displays were built in.

Figure 1: Growing amount and size of electronic displays in passenger cars (Isele 2004)

One of the main benefits of an electronic display compared to a conventional analogue interface is the fact, that it is freely addressable. In order to be able to display 10 warning messages an analogue cluster instrument must have all ten warning symbols engraved in the panel. The symbol corresponding to a special warning message gets highlighted the moment the driver needs to be warned. With an electronic display the space for one warning symbol would suffice to be able to display any warning symbol when necessary and in the meantime the space could even be used to display other information. The two effects saving space and the possibility of pre-filtering the amount of information conveyed to the driver (thus reducing the driver's cognitive workload) led to the increased use of

electronic displays as user interface to more and more functions. For example the BMW i-drive concept would not have been possible without a high performance electronic display. The i-drive system pools all comfort and entertainment functions into a shared Man-Machine-Interface (MMI). Using a single interface for the radio and CD player, the climate control, the navigation system, the hands free phone or other information from the bordcomputer, would occupy a lot of space. Instead these entire devices share the central information display (CID) placed in the middle of the cockpit and navigation through the menus as well as selection of functions can be done single handed by a single knob within the drivers reach.

Even classical analogue instruments like the cluster instrument incorporating speedometer and tachometer include more and more electronic displays. First only the space between the speedometer and tachometer was used for electronic displays displaying information like mileage or fuel consumption. In 2002 BMW introduced an instrument cluster filling the inner circles of the tachometer and speedometer with an LCD while still having mechanical pointers and fixed numbers for speed and rpm. In future even a fully digital cluster instrument might be possible.

Highly interesting display technologies for use in the dashboard are bi-stable reflective display technologies, i.e. reflective displays which are stable (show an image) in the on and in the off-state, power is only needed to change the displayed image. The key benefit of bi-stability is low power consumption and the possibility to show an image in the off-state. The latter feature is of special interest to automotive designers aiming at making the dashboard look interesting in the

showroom even when the car is turned off. With a bi-stable reflective display this could be done because these displays can show an image in the off state instead of a black screen conventional displays show when switched off. Furthermore reflective displays do not have to compete against daylight.

As electronic displays become more and more important physical Man-Machine-Interface devices they attract growing attention from the driver. As a result display quality becomes an issue. The BMW slogan “Sheer driving pleasure” means that driving a BMW is supposed to be fun and enjoyable. The driver’s perception of his driving experience is placed into the focus of attention. Of course, the driving experience does not stop at the driving impression but includes other perceptual aspects like for example perceived comfort or design. For the automotive display this means the ultimate benchmark for display quality is the driver’s perception of the display quality.

The aim of this thesis is to link perceived image quality to physical display parameters. This is done for the physical display parameters black and white luminance, contrast and colour gamut.

The thesis is structured into 8 chapters. While chapter 1 is this introduction chapter 2 provides background material. It deals with basics of human perception as well as basic display parameters and special challenges posed to the display and observer by the automotive environment. Chapter 3 introduces three image quality models relevant to automotive applications from the group of readability models as well as the framework of image quality modelling by Engeldrum (2000), which is

the basis for the experiments performed in chapter 7. As the image-quality-circle introduced by Engeldrum used psychometric scaling to model image quality, an introduction to psychometric scaling methods is given in chapter 4. It has already been mentioned that reflective displays are an interesting technology for automotive applications, because of their good sunlight-readability and the new design opportunities bi-stable electronic displays could offer. For this reason a description of reflective display technologies is included in chapter 5. One of the key challenges to automotive displays are the ambient lighting conditions the display will have to face in the automotive environment. In chapter 6 possibilities to simulate daylight either electronically or by mock-ups as used in architecture or colour matching are summarized. Chapter 7 is the experimental part where the influence of luminance, contrast and colour gamut on perceived image quality is investigated. The thesis finishes with conclusions and suggestions on further work in chapter 8.

2 Background Material

As this thesis deals with perceived image quality of automotive displays the first part of this chapter gives an introduction to basics of human perception. Already from 1920 on the pioneers of psychophysics have demonstrated that human perception can indeed be measured and algorithms linking physical parameters to perceptions exist.

The second part of this chapter deals with the physical parameters of automotive displays. For perceived image quality the optical parameters are the most relevant while lighting conditions influence the display as well as the driver.

2.1 Basics of human perception

As we are looking for a way to quantify perception it is necessary to have a look at some basics of human perception and corresponding basic measurement techniques first. Goldstein (2008) describes perception as an iterative process which can be represented by a circle:

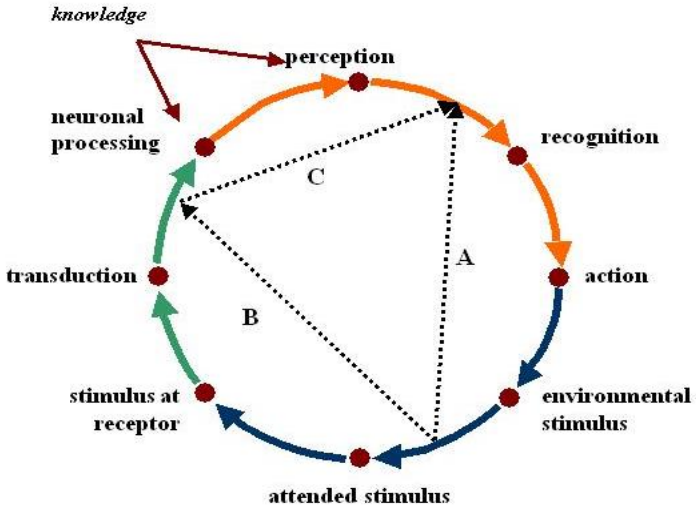

Figure 2: The perception process after Goldstein (Goldstein 2008, p4)

When following the human signal processing bottom-up, out of the wealth of stimuli present in our environment, perception starts with a special physical stimulus we pay attention to: the **attended stimulus**. This can be a sound, a smell, a touch, a taste or light entering the eye. The first processing stage of this stimulus occurs at the corresponding **receptor** be it the ear, the nose, the skin, the mouth or the eyes. These receptors transform the physical stimuli into electrical signals in a process called **transduction**. The generated electrical signals are transmitted via a complicated and not yet fully understood net of neuronal connections to the brain while some **neuronal processing** takes place already. In the brain the signal passes through several processing stages. First the signal is **detected**, then it is **recognised** the decision for an appropriate **reaction** to the stimulus is taken and performed. For the processing stages taking place in the brain former knowledge has an influence leading to a certain degree of top-down processing. For example it is harder to make out single words in an unknown language, but with growing vocabulary and hearing experience we are able to identify individual words we have learned and have less and less the impression of a continuous stream of strange sounds.

The arrows A and B indicate two approaches to the study of perception. Physiological research tries to understand and model the inner processes of perception represented by the arrows B and C. Arrow B leading from attended stimuli to activated brain regions represents the physiological reaction while arrow C follows the processing stages in the human brain from the physiological reaction to perception, recognition and action. The human physiology sets the scene for which stimuli can be perceived.

In psychophysics the relationship between perception (psycho-) and the physical stimulus (-physics), represented by arrow A, is studied.

In order to measure perception it has to be assessed in some way. There are several levels of assessing a perception which will be illustrated by the example of the perception of having a cocktail in a bar:

Description: The naming of attributes of a stimulus: “The cocktail the waiter puts in front of me is greenish. It tastes sweet, creamy and like menthe.”

Recognition: To classify the stimulus: “This cocktail is a grasshopper.”

Detection: To become aware of a difficult to detect aspect of a stimulus. “They must have added a hint of pistachio.”

Perceiving values: To be aware of the size or intensity of a stimulus **“**My cocktail is a third sweeter than yours.”

Search: Finding a stimulus among a number of other stimuli. “The girl on the third table on the left is drinking a grasshopper, too.”

Important early contributors to the research domain of psychophysics were Fechner and Weber. They were the first to prove that perception can be quantified. Fechner introduced three classical measurement methods to detect the **absolute threshold** of perception (Goldstein 2008, p13). In the **threshold method** the physical stimulus is varied in discrete ascending or descending steps and the subject is asked whether or not it perceives the stimulus. The threshold then is the average between the detected and not detected stimuli. In the **production method** the subject can directly control the stimuli continuously and is asked to adjust the stimulus to a value it can just perceive. The most accurate method is the

continuous method where the subject is presented several times in random with stimuli around the threshold and is asked whether or not it perceives the presented stimulus. When plotting the number of times a stimulus of each value was perceived in percent an s-shaped graph the so called psychometric curve is the result.

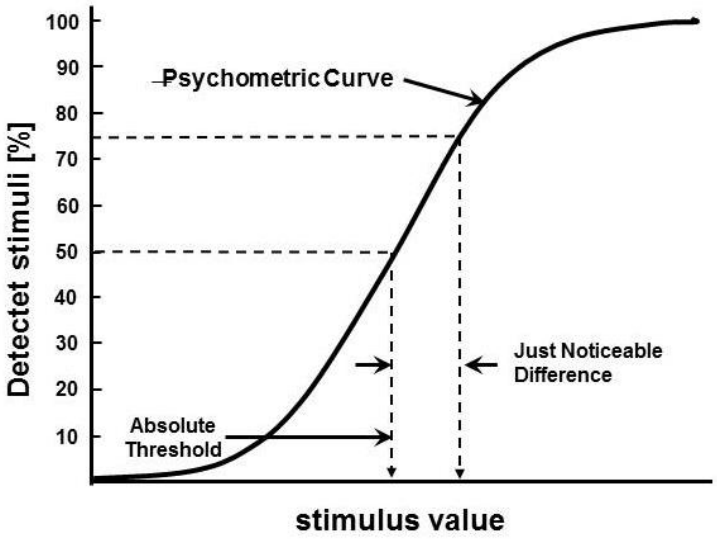

Figure 3: The psychometric curve

A common definition of the absolute threshold is the stimulus value at which 50% of the representations are perceived. Since Fechner a number of further and for some applications more effective psychophysical experimental methods were established, some of which will be described in detail in chapter 4.

Weber went a step further and introduced **relative thresholds**. He found out, that the ratio between a stimulus and the increase or decrease in the stimulus necessary to detect a difference is constant. For example to detect a difference in weight relative to a 500g box of candies, 10g of candies have to be added to

perceive the difference. For a big box of 1000g of candies an extra 20g of candies is necessary to perceive a difference in weight. The ratio in both cases is 0.02 i.e. constant. The corresponding law often called Weber-Fechner law is:

$$\Delta s / s = c$$

Equation 1: Weber-Law (Goldstein 2008, p15)

In equation 1 s is the stimulus, Δs is the increase in stimulus leading to perception and c is the Weber constant. As seen before the Weber constant for holding a weight is 0.02. For other perceptions different Weber constants apply like 0.08 for brightness, 0.04 for loudness or 0.01 for pain induced by electro shocks (Goldstein 2007, p15).

However, when one intends to quantify a perception threshold values are of limited help. They do not answer questions like whether doubling or tripling stimulus intensity leads to a perception twice or thrice as strong.

In order to **scale perception** Stevens introduced **magnitude estimation** and **magnitude production methods** to scale perception. In magnitude estimation the subject is presented with a standard stimulus which has an assigned number. The task is to assign a number to a test stimulus which relates to the perceived intensity of the test stimulus relative to the standard stimulus. The use of a standard stimulus is a help to the subject giving him an anchor point, which does not necessarily have to be used. Magnitude estimation experiments can also be

performed by asking the subject directly to assign numbers to stimuli relating to the perceived stimulus intensity.

In magnitude production the subject is asked to adjust a stimulus to for example twice the intensity of a standard stimulus. Stevens did these experiments for various perceptions like brightness, heaviness of lifted weights or pain induced by electro shocks, to name just a few.

The finding of Stevens represented by **Stevens Power Law** is that for all perceptions he examined there exists a common relationship between the stimulus intensity and the perceived intensity:

$$\Psi(I) = kI^{\alpha}$$

Equation 2: Stevens Power Law (Stevens 1959)

Where Ψ(I) is the perceived intensity of the stimulus, I is the magnitude of the physical stimulus, the exponent α depends on the type of stimulation and k is a proportional constant. For example the exponent α for length perception is 1 which means that there is a linear relationship between the physical length of a line and our perception of the line's length. Pain induced by electroshocks has a rather high exponent of 3.5 illustrating the warning function of pain. (Goldstein 2008, p15)

Brightness of a 5° viewing area in the dark has an exponent of 0.33 reflecting sinking sensitivity at higher brightness levels (Goldstein 2008, p15). A similar

relationship can be found in the Lightness formula incorporated in the CIE 1976 L*u*v* colour space which is supposed to be perceptually uniform:

$$L^* = 116\left(\frac{Y}{Y_n}\right)^{1/3} - 16$$

Equation 3: CIE Lightness formula (CIE 2004)

Both the CIE L*u*v* and L*a*b* colour spaces, which will be explained in more detail later on, were initially intended for modelling object colours. Hence the quantity L* represents lightness i.e. the extent to which a stimulus appears to emit more or less light than a similarly illuminated stimulus that appears white. Therefore Y_n is specified as the white object stimulus and Y the CIE 1931 Y primary of the measured object. Usually Y_n is taken to be equal to 100 representing the luminance factor of a perfect reflecting diffuser.

Even though the CIE formula is intended for lightness of reflecting objects it is in practice used for self-luminous objects like displays as well. An interesting question is whether it predicts the brightness perception of self-luminous objects even better since these have an exponent α of 0.33 i.e. 1/3 in Stevens power law while for lightness perception resulting of reflectance from grey paper an exponent of 1.2 is used.

2.2 The automotive display

Let's apply Goldstein's circle of the human perception process to the driver's perception of a warning message lighting up in his cluster instrument.

Figure 4: Environmental stimuli at the driver working place (BMW Group 2008)

The environmental stimuli are given by the stimuli emitted from the complete scene the driver is seeing like for example in figure 4. The lighting up warning symbol changes the light pattern on the retina i.e. the stimulus at the receptor. The visual receptors transform the light pattern into electrical signals which are pre-processed by the human visual system's neuronal network and transmitted to the corresponding areas of the human brain. The driver recognises that an additional light appearing in the area of the cluster instrument must be an important warning message and performs the action of focussing on the warning messages which becomes the attended stimulus. The driver's viewing behaviour while driving will be looked at in more detail later on. The image of the warning symbol on the retina again is transformed into electrical signals wired through the neuronal network to the brain processed there and recognised as general warning message whereupon the driver decides to stop the car and redirects his gaze on the street in order to keep lookout for the next parking possibility. This example shows that perception is a continuous iterative process where former knowledge or so called cognitive aspects like expectation play a role.

This part introduced human perception as an iterative process which can be represented by a circle (Goldstein 2008). The two main approaches to the study of perception, physiological research and psychophysics, target the relationship between different processing stages in the perception process. This thesis belongs to the field of psychophysics, linking physical stimuli directly to perceptions. Examples from the research of Weber, Fechner and Stevens were presented, demonstrating that perceptions can indeed be quantified. Finally this part was finished by an example of the perception process in an automotive setting in the case of a warning signal lighting up in the cluster instrument. This part of the background chapter dealt with basics of perception. As this thesis is about perceived image quality of automotive displays, the next part this chapter will describe the physical parameters of an automotive display as well as influences which are special to the automotive environment.

2.2.1 Mechanical and thermal display parameters

The automotive environment has not only a challenging dynamic in the ambient illumination. The temperature range the display is faced with is equally demanding. Electrical or electronic assemblies in the passenger compartment which are subject to direct sunlight are supposed to withstand temperatures of up to 105 °C. The other extreme is wintertime when temperatures in the parked car are going down and the display is supposed to be operational at temperatures as low as –40 °C (BMW 2007a). A display must be able to survive these temperatures in the off state and be able to start up from these temperatures.

Mechanical display parameters like total display size and diagonal (also called form factor) as well as pixel size and pitch or overall weight of the display influence usability of a display for certain applications. A good example for the importance of form factor is the crowding out of CRT computer monitors and TV-sets by flat-panel LCDs. The LCDs don't necessarily provide better image quality, but it is sufficient and significantly less space is required for the display. In automotive applications where a great number of devices compete heavily over construction space a good form factor is of utmost importance. The switch from CRT to LCD within BMW cars happened already in the year 1994 with the introduction of the 4.8" bordcomputer.

Display weight is more important for portable applications like PDAs or mobile phones as well as laptop computers. Pixel size and pitch however directly influence display perception. As the human eye can't resolve details smaller than 1 arc minute this is the maximum pixel pitch possible, if the observer shall not resolve the pixel structure of the display. In automotive applications the typical viewing distance is around 80 cm. That gives a desirable pixel pitch of around 0.233 mm.

Overall, a certain degree of robustness is necessary be it to vibration, touch or abrupt acceleration and deceleration.

For most display technologies meeting the thermal requirements is the hardest part. OLEDs for example downrightly melt at the high automotive temperatures and are extremely sensitive to humidity. LCDs on the other hand have problems at freezing temperatures where the liquid crystals tend to stick. A great number of display technologies exhibit a temperature dependence of their optical properties.

For example in cholesteric LCDs the pitch of the bragg grating, which is responsible for the reflected colour, is temperature dependent.

2.2.2 Optical Display Parameters

Optical display parameters basically are all the parameters by which the light spectrum emitted of and/or reflected from a display can be described. Be it in terms of basic photometric quantities, colour coordinates or spectral and angular reflectance properties.

2.2.2.1 Basic Photometry

In contrast to radiometry where electromagnetic radiation is taken as it is, in photometry the radiation is weighted with the sensitivity function of the human eye. This function is called the V(λ) function for photopic vision and is supposed to correlate to the visual sensation of a standard human observer exposed to the same radiation. Photopic vision is also referred to as daylight vision and therefore most important for automotive display applications. Another sensitivity function for scotopic or night vision is basically of the same form but shifted towards shorter wavelengths. However, the scotopic region generally is not reached in automotive applications. When driving at night there are enough bright sources around to lift the driver's sensitivity to the so called mesopic region. The mesopic region is a transition region between the photopic region where only cones contribute and the scotopic region where only rods contribute. The mechanisms of mesopic adaptation are still under research (Eloholma et. al. 2005, Walkey et. al. 2006).

For display applications the most important photometric quantities are the luminance L of the display and the illuminance E falling on the display due to

ambient illumination. As these quantities are based on the luminous intensity respectively luminous flux emitted from or falling on an area all four basic photometric quantities are described briefly:

Radiant flux and luminous flux

The radiant flux is the energy radiated by a source per unit of time:

$$\Phi_e = \frac{dQ}{dt}$$

Equation 4: Radiant Flux (ISO 23539)

Where Q is energy t is time and ϕ_e is the radiant flux measured in watts. The little e stands for energetic. The photometric quantity luminous flux is obtained by weighting the radiometric quantity radiant flux with the V(λ) function:

$$\Phi_v = K_m \int_{380}^{780} \Phi_{e,\lambda} V(\lambda) d\lambda$$

Equation 5: Luminous Flux (ISO 23539)

The visual (subscript v) quantity luminous flux has been given the unit lumen [lm]. K_m is the maximum spectral efficacy of radiation for photopic vision. The value of K_m is 683 [lm/W].

Radiant intensity and luminous intensity

Radiant intensity is defined as the energy flux per unit solid angle in a given direction from the source. The corresponding photometric unit is the candela [cd]. As the first photometric standards were candles the candela happens to be the photometric base unit (SI unit). One candela is defined as the luminous intensity in a given direction of a source that emits monochromatic radiation of 540* 10^{12} Hz i.e. 555nm and has a radiant intensity in that direction of 1/683 W/sr.

$$I = \frac{d\Phi}{d\Omega}$$

Equation 6: Radiant and luminous intensity (ISO 23539)

Irradiance and illuminance

Illuminance E is defined as luminous flux per detector area and has been given the unit name lux (=cd/m²).

$$E = \frac{d\Phi_S}{dA_D}$$

Equation 7: Irradiance and Illuminance (ISO 23539)

The subscript S stands for source and the subscript D for detector.

Radiance and luminance

Luminance L is defined as luminous intensity I in candela [cd] emanating from a plane per square meter.

$$L = \frac{I}{A_S}$$

Equation 8: Radiance and Luminance (ISO 23539)

Luminance is the photometric quantity used for measuring displays.

2.2.2.2 Basic Colourimetry

The still most widely applied system for unambiguous description, calculation and measurement of colours is the CIE1931 colour space (CIE 2004). The two-dimensional representation of the CIE1931 colour space, the chromaticity chart, is its most commonly seen graphical representation. The chromaticity coordinates x and y are derived by normalisation to the CIE 1931 tristimulus values X, Y, Z:

$$x = \frac{X}{X+Y+Z} \qquad y = \frac{Y}{X+Y+Z} \qquad z = \frac{Z}{X+Y+Z} \qquad x+y+z = 1$$

Equation 9: CIE1931 chromaticity coordinates (CIE 2004)

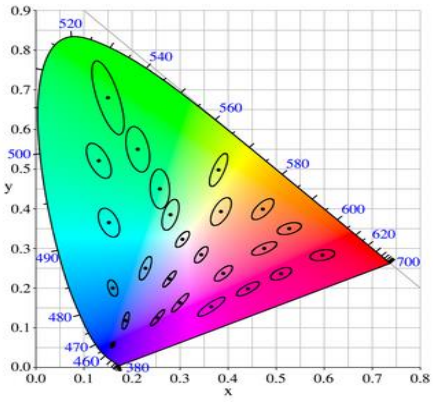

Figure 5: CIE1931 chromaticity diagram with 10 fold MacAdam ellipses

The black ellipses in the chromaticity diagram are the so-called MacAdam ellipses. MacAdam (1942, 1943) established colour difference thresholds for a large number of colour loci within the CIE1931 colour space. In a perceptually uniform colour space the loci of just noticeable colour differences would form a perfect ball around the reference colour. In a perceptually uniform two-dimensional representation the iso-threshold lines would form perfect concentric circles around the reference colour. In CIE1931 however, the threshold lines do not form circles, but ellipses of varying shape and size. That means although CIE1931 is based on psychophysical experiments and highly practical to work with, it is not perceptually uniform.

The desire to obtain a perceptually relevant figure for colour difference led to the development of the two-dimensional CIE1976 uniform colour scale (UCS) and the CIE1976 L*a*b* (short CIELAB) and CIE1976 L*u*v* (short CIELUV) colour spaces. The UCS basically is a slightly distorted CIE1931 chromaticity diagram:

$$u' = \frac{4x}{x + 15y + 3z}$$

$$v' = \frac{9x}{x + 15y + 3z}$$

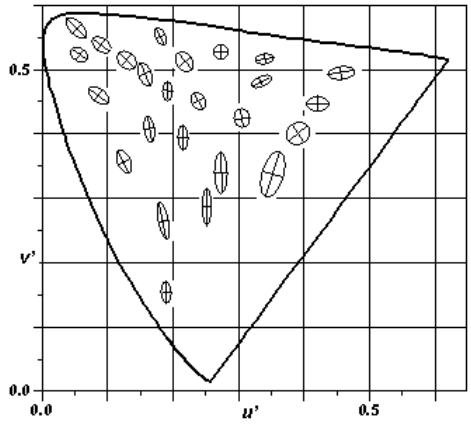

Figure 6: CIE1976 UCS diagram with 10 fold MacAdam elipses (Judd and Wyszecki 1975)

The first models taking lightness and chromatic adaptation into account were the CIE1976 colour spaces CIE L*a*b* (short CIELAB) and CIE L*u*v* (CIELUV). These models take the inherent red/green, yellow/blue and light/dark processes of the human visual system into account as well as a basic chromatic adaptation transformation. Both colour spaces feature three dimensions which are supposed to correlate with perceived lightness, chroma and hue of a stimulus. The main difference is the chromatic adaptation function applied. While CIELAB coordinates are established by normalising to the white point, in CIELUV accounting for chromatic adaptation is done by a shift:

$L^* = 116\left(\frac{Y}{Y_n}\right)^{1/3} - 16$	$L^* = 116\left(\frac{Y}{Y_n}\right)^{1/3} - 16$
$a^* = 500\left[\left(\frac{X}{Xn}\right)^{1/3} - \left(\frac{Y}{Y_n}\right)^{1/3}\right]$	$u^* = 13L^*(u' - u'_n)$
$b^* = 200\left[\left(\frac{Y}{Y_n}\right)^{1/3} - \left(\frac{Z}{Z_n}\right)^{1/3}\right]$	$v^* = 13L^*(v' - v'_n)$
$C_{a,b}{}^* = \sqrt{(a^{*2} + b^{*2})}$	$C^*_{uv} = \sqrt{(u^{*2} + v^{*2})}$
$h_{a,b} = \arctan(b^*/a^*)$	$h_{uv} = \arctan(v^*/u^*)$
$\Delta E^*_{ab} = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2}$	$\Delta E^*_{uv} = \sqrt{(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2}$

Table 1 CIELAB and CIELUV Colour Space Formulae (after CIE 2004)

Of both 1976 colour spaces CIELUV is more popular within the display industry. The main reason is that additive mixtures plot as straight lines in CIELUV while they lead to curved lines in CIELAB. Furthermore CIELUV provides a chromaticity

diagram in which colours can be plotted independently of their L* values and in the associated 1976 UCS diagram the difficulty of providing a relevant white object stimulus can be avoided completely.

2.2.2.3 Display reflectance

Generally display reflectance can be divided into three types of reflectance with fundamentally different appearances: Diffuse lambertian reflectance, specular reflectance and haze. The figure 7 shows what a dark display with a front screen exhibiting only the named kind of reflectance will look like (VESA 2001):

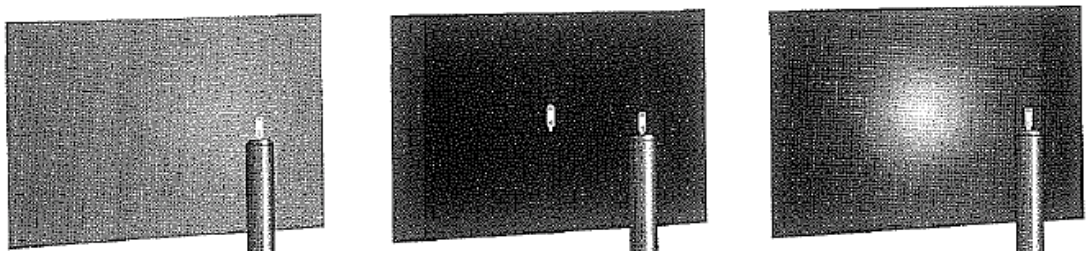

Figure 7: Appearance of a pure lambertian or specular or haze component (VESA 2001)

Pure specular reflectance produces a distinct image of the light source on the material surface while pointing a point source at a lambertian material still results in a quite uniform reflection over the whole material. A haze component manifests itself as a fussy ball around the distinct image of the light source.

To fully characterize the reflectance distribution of a material the angular distribution of reflected light depending on the illumination angle is needed. Such a function is called bi-directional reflection distribution function (BRDF). For materials exhibiting spatially varying reflectance properties like textured cloth an even more complicated bi-directional texture distribution function (BTDF) may be

necessary. These functions are used as “optical textures” in lighting simulation programs.

The problem is, that even though the name bi-directional reflectance distribution function sounds like just typing in a formula to characterize the optical properties of the material in question, in reality it implies a look-up table with a lot of measurement values. Data for a complete BRDF implies sampling the reflectance over the half sphere on top of a material for all possible incidence directions. As this is definitely a lot of data, which is time consuming to gather and demands a lot of computational power in the simulation, in most cases simplifications are used. One possibility is to elicit the BRDF only for one incidence angle. Another possibility used for materials with rotational symmetry in their reflection properties is to use just a vertical section of the reflected cone. As can be imagined, the amount of data to be gathered for a BTDF, where the sampling is not done for a point but for an area, is even higher.

The diagram below shows the cross-section of a BRDF graph. The peak at the incidence angle of the light source represents the specular component of the reflection and the uniform reflectance base of the reflection represents the lambertian component. What is often causing trouble is the haze component.

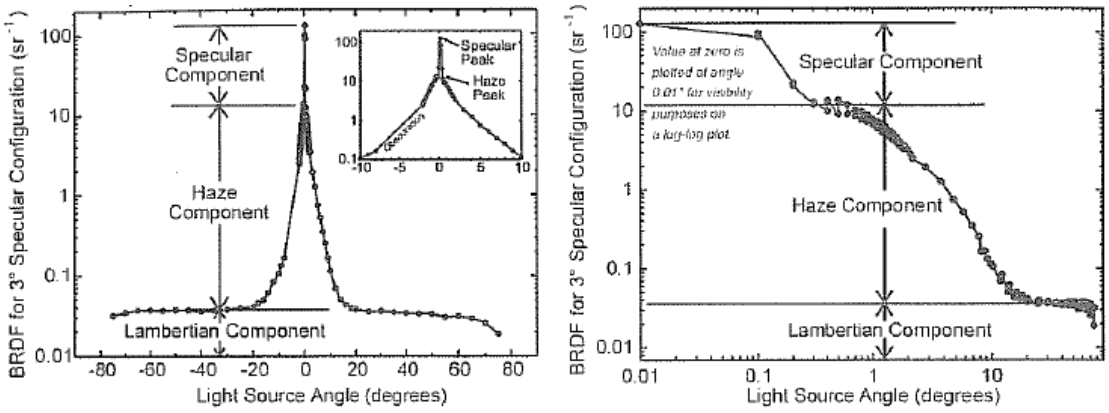

Figure 8: BRDF of a display with lambertian, haze and specular component (Kelly 2002)

In the graphs in figure 8, the haze component is the area around the peak of the specular reflection which has significantly higher reflectance than the lambertian part. Many surface treatments which scatter some of the specular light energy induce haze. The difficulty is that the haze region is directly around the very narrow specular reflectance cone. Depending on measurement aperture or size of a light trap around the specular direction more or less of the haze component is detected thus making especially displays with a haze or even worse a nontrivial haze component especially sensitive to measurement geometry (Kelly 2002).

In practice one still only accounts for the lambertian (diffuse uniform distribution) part and the specular part given by the following rather simple equations:

$$L = \frac{\rho}{\pi} E$$

Equation 10: Luminance /Illuminance Relationship Lambertian Diffuser (Kelly 2002)

$$L = \rho_s L_s$$

Equation 11: Specular reflection (Kelly 2002)

In equation 11 L is the reflected luminance, E the illuminance, ρ the diffuse reflectance, ρ_s the specular reflectance and L_s the source luminance.

2.2.3 Lighting conditions

As can already be seen in the discussion of a display's reflection properties, the stimuli of the display are to some extent dependent on or superimposed by environmental stimuli. Furthermore, perception of the stimuli is to some extent influenced by environmental conditions. One quick example is that object colours can't be perceived in the dark. Furthermore environmental conditions can have a direct influence on the display's physical image parameters. The most important environmental condition for display appearance is ambient lighting.

ISO (2001) establishes two automotive daylight scenarios, diffuse illumination of 3 klux and direct illumination with 45 klux under 45° to the viewing direction or under the car's individual glare angle. The 3 klux diffuse scenario is to represent standard daylight conditions while the 45 klux under glare angle is supposed to mimic a worst case scenario.

For measurement purposes one can deal only with the detrimental effect on the display. That means for display measurements it is sufficient to keep the set-up geometries as simple as possible and illuminate only the display. As soon as the

driver’s perception of the display is to be evaluated, the driver’s adaptation luminance has to be accounted for. Tables 2-4 show the concrete values for lighting scenarios used in readability models (described in detail in section 3). All of these scenarios do account in some form for the observer’s adaptation luminance. A set-up for psychometric measurements of display appearance should ideally provide a controlled display illuminance as well as a controlled driver adaptation luminance.

Scenario	Display direct illuminance [lux]	Display diffuse illuminance [lux]
Direct Sun	45,000	
Diffuse Day		3,000

Table 2: Lighting Scenarios ISO/Dis 15008 (ISO 2001)

Scenario	Display direct illuminance [lux]	Display ambient illuminance[lux]	Forward scene [lux]
High Ambient Sun Rear	100,000		15,000 Visor down
High Ambient Sun Front		8,000	110,000 Visor down
Low Ambient Dusk		1,500	200

Table 3: Lighting Scenarios PJND Modell (BAE SYSTEMS 2001)

Scenario	Diffuse [lux]	Specular [lux}	Glare [cd/m2]
Sun Rear	100,000	20,000	75
Sun Forward	10,000	70,000	350
High Ambient	20,000	20,000	75

Table 4: Daylight Scenarios after Sharpe et. al (2006)

Condition	Ambient Illumination [lux]	FFOV [cd/m²]
Office (Indoor)	500	100
Conference room	10	171 (bright uv-gap)
Direct Sun Light (outdoor)	100,000	34,260
Car – night	1	0.3
Car – day / lo-ambient	3,000	3,000 (clear blue sky)
Car – day / med-ambient	20,000	10,000 (haze)
Car – day / high-ambient	60,000	34,260 (sunlit clouds)
Car – day /hi - FFOV	3,000	34,260 (sunlit clouds)

Table 5:Time to visibility (TTV) model (Silverstein 1996):
Ambient illumination and field of foreward view (FFOV) Conditions

2.2.4 Driver viewing behaviour

This thesis deals with the perception of displays within a car. In contrast to most other display applications looking at the display is NOT main task in an automotive scenario. The driver's main task is to drive safely. The cluster instrument which displays information like the current velocity and engine speed as well as warning messages directly assists the driver in the driving task. Other displays like the Central Information Display (CID) offer additional information like navigation maps as well as entertainment and comfort functions like radio/CD or climate control. According to its relevance for the driving task the cluster instrument is placed near the driver's main viewing direction i.e. directly in front of the driver and 20° down. The CID, which shall not distract the driver from driving safely, is placed outside the main viewing direction at 60° to the right of the driver and 20° down.

The first consequence is that viewing geometries of automotive displays differ significantly from other display applications where the viewer tends to be directly in front of the display and look at it perpendicularly or slightly down.

The second main difference to display applications is that the driver does not look continually at the display, but only directs short glances at it. For this reason the second consequence is, that information has to be accessible very fast. In a publication by the U.S Department of Transportation (DOT) on "Suggested Procedures and Acceptance Limits for Assessing the Safety and Ease of Use of Driver Information Systems " (DOT 1995) figure 8 is presented, based on research by Zwahlen et. al. Figure 9 shows acceptable and unacceptable durations of glances at a display or control in the car via numbers of glances necessary to retrieve information from a display or operate a control. The acceptable duration of

a glance gets shorter with the number of glances necessary to obtain information from the display. While a glance duration of 1.2 s is totally acceptable, if one look at the display is enough to obtain the required information only 0.9 s are perfectly acceptable if three glances are necessary. One glance which takes between 1.2 s and 2 s falls into a grey area acting as a soft border between acceptable and unacceptable operation. A display's usability inside a car is definitely unacceptable, if a driver needs to look longer than 2 s at one glance, 1.5 s at 4 glances or more often than 4 times at the display in order perform the desired task. Easy readability of displayed information is therefore the key quality aspect of many image quality models related to automotive or avionics applications.

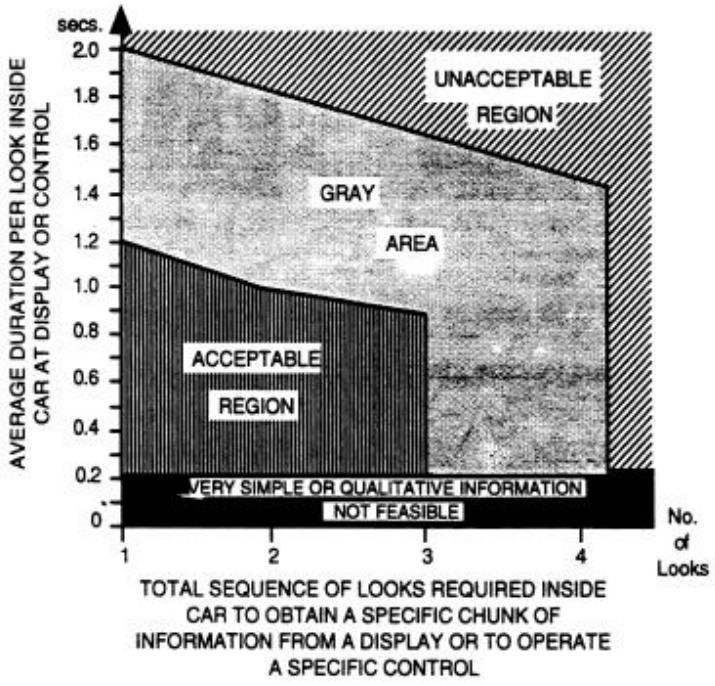

Figure 9: DOT eye fixation recommendations for driver information systems (Zwahlen in DOT1995)

2.2.5 Summary automotive display

This section on the automotive display covered physical parameters of displays and highlighted special challenges the automotive environment poses for electronic displays. Out of the described physical parameters the optical display parameters are the relevant stimuli, which will be used as input for psychometric scaling of the observer's visual perception of the display. The optical image parameters used in the experiments will be luminance and CIE xy colour coordinates. Illuminance will be reported in case displays are illuminated in an experimental setup. Even though reflection properties of the display and glance duration are expected to have an impact on perceived image quality they will not be investigated in this thesis.

3 Image Quality Models

In the background chapter basics of perception and physical parameters of automotive displays were covered. As this thesis is titled “Modelling Image Quality for automotive Display Technologies” image quality models relevant to automotive applications are the subject of this chapter. In the first part of this chapter three image quality models, which are of value for the automotive industry because of their ability to predict readability of information on the display, are described in detail. Each of these models has its individual strengths in predicting the readability of information the display under varying lighting conditions or even for different font sizes. For an automotive display it is of utmost importance, that safety relevant information is readable under all environmental conditions, therefore it is no small task to be able to predict readability. However, even a perfectly readable display is not necessarily regarded as a high quality display. Once readability is assured, the next step is to rate the perceived quality of the display. To this end the framework of the image quality circle by Engeldrum (2000) is introduced in the final part of this chapter.

Researchers from various backgrounds have developed models of human visual perception. One way to group them is by the purpose of these models. Classical Image Quality Models mostly aim at predicting the visibility of compression or other undesirable artefacts within the image. This is usually done by comparing an original image and a processed image through a filter incorporating spatial sensitivities and masking effects associated to a so called standard observer. These models are not especially helpful in determining image quality of automotive displays as they are more targeted on evaluating image processing algorithms

than evaluating the capabilities of the display. Furthermore these models always require a reference.

Another group of vision models are especially designed for sunlit environments or even directly dedicated to the automotive environment. These models use knowledge about the human visual system to determine whether information in a given environment is visible or readable. These models will be looked at in greater detail, as they are very useful in ensuring that the automotive display is fit for its basic purpose to deliver information. However, ensuring that a display does its job may keep it from definitely being a bad display but does not necessarily make it a good one.

Image quality is not only given by the absence of artefacts and ensured functionality. It is necessary to identify the aspects making a good display and to quantify them. To achieve this goal a psychophysical approach will be used as described in the last part of this chapter.

3.1 Readability/Visibility Models

Readability models belong to the group of performance models. They don’t aim to reproduce exact appearances or model detailed processing steps. Readability models concentrate on whether or not a viewer is able to read given information under given conditions. Most of these models are based on empirical studies and an underlying model of human vision.

3.1.1 Perceived Just Noticeable Difference Model (PJND-Model)

The Perceived Just Noticeable Difference (PJND) model was developed within the avionics industry (BAE Systems 2001). This model is interesting for automotive applications for several reasons:

- The impact of varying lighting scenarios on the luminance and colour contrast of a display is calculated
- A figure of merit for contrast perception which is independent from the actual luminance levels is provided
- Luminance/illumination level is given
- The PJND is not a digital visible/not visible decision but a scalable figure of merit
- For avionic applications a scale assigning a necessary number of PJNDs to specific tasks already exists

Like cars aircraft are operated outdoors. This means built in displays and their users' experience dramatically changing lighting conditions. Within an aircraft the ambient illumination conditions range from complete darkness at night-time flights to over 100,000 lux of daylight when flying over the cloud cover (BAE Systems 2001). For the PJND model the following three worst-case lighting scenarios were considered most important:

High Ambient Sun Rear Scenario (BAE Systems 2001)

This scenario is especially detrimental to emissive displays, because the reflected sunlight reduces the display's contrast and saturation. In detail this scenario is defined by the aircraft flying straight and level at 30,000 feet above a full cloud cover, the solar disk being at a low elevation angle of 30° to the rear of the aircraft, the display being submitted to direct sunlight of up to 100,000 lux with specular reflections from the display and the general cockpit area predominating the scene. The aircrew is supposed to be wearing a tinted visor while looking at a diffused forward scene approaching 15,000 lux.

High Ambient Sun Forward Scenario (BAE Systems 2001)

In the sun forward scenario the low sun directly shines into the pilot's eyes thus inducing veiling glare. In detail the orientation of the aircraft is level at 30,000 feet over a full cloud cover like in the sun rear scenario. The solar disc is even lower at 15° elevation forward of the aircraft. The display is in the shadows illuminated by diffused skylight approaching 8,000 lux. The display is mainly influenced by specular reflections from the pilot's flying suit. The aircrew helmet tinted visor is down and the forward scene is dominated by the solar disk, approaching 110,000 lux.

Dusk/ Dawn Scenario (BAE Systems 2001)

The third condition is especially detrimental to reflective displays. At dusk/dawn with the sun in front of the aircraft and close to the horizon the cockpit is in hard shadow while the pilot is subjected to solar glare. This means the pilot's perception capabilities are degraded while only little luminance is reflected by reflective

controls. In detail the solar disc is in front of the aircraft and close to the horizon line producing 1,500 lux of ambient light from the forward direction. The sky hemisphere is clear and of low luminosity approaching 200 lux. The cockpit lighting control 'Night' is selected and set to maximum and the aircrew helmet visor is up.

EmF	Emission Foreground
EmB	Emission Background
DrF	Diffuse Reflection Foreground
DrB	Diffuse Reflection Background
SrF	Specular Reflection Foreground
SrB	Specular Reflection Background
SAR	Specular Ambient Reflections
DAR	Diffuse Ambient Reflections
GAR	Glare Ambient Reflections

Table 6: Abbreviations used in Lighting Scenarios (BAE Systems 2001)

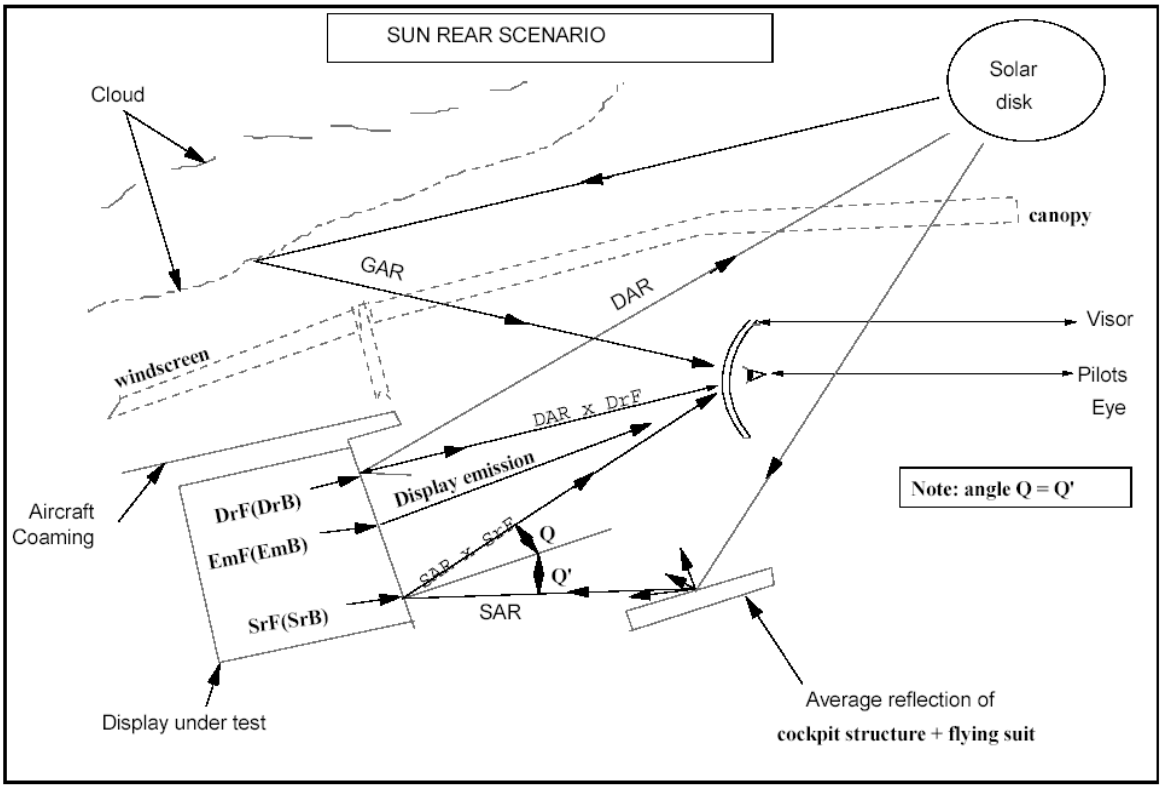

Figure 10: High Ambient Sun Rear Scenario (BAE Systems 2001)

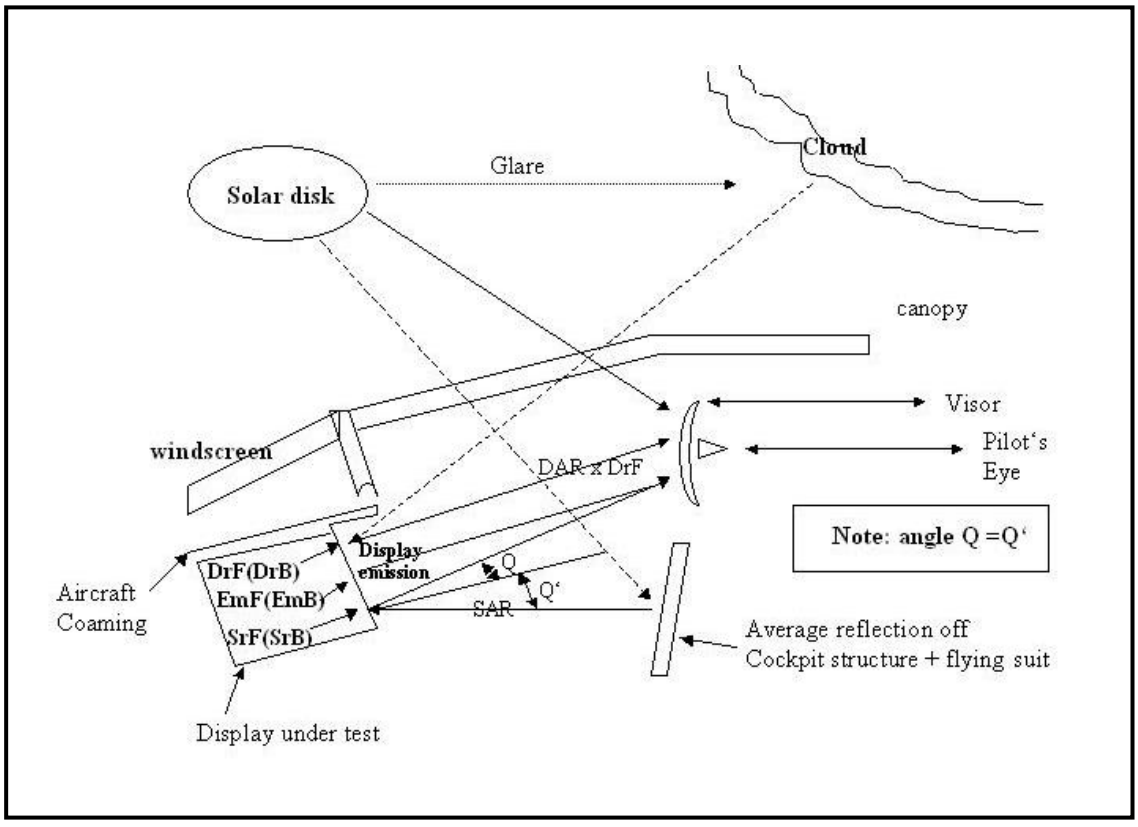

Figure 11: High Ambient Sun Forward Scenario (BAE Systems 2001)

The approach of defining critical environments and calculating their influence on display performance is highly relevant to automotive applications as well. Due to safety reasons automotive displays must be readable under all possible conditions and the described scenarios are a good example for critical conditions which will occur on the ground as well.

On the ground the corresponding illuminance values might differ and will vary with geographical position. Unfortunately, drivers are much less predictable than pilots in terms of reflectance from their suits, transmission of their sunshades or level to which they switch their internal-lighting control. Viewing-geometries are different to those in aircraft. Depending on the relative position of the automotive display, solar disc elevation levels leading to glare or frontal illumination of the display may vary significantly. The huge variety of interior materials and colours may lead to a considerable variation of specular ambient reflections having impact on the display. However, if one bears in mind that for safety reasons readability under all possible viewing conditions has to be guaranteed the variations can be reduced to worst case scenarios. Solar data for the whole world is available and could be used to create worst case automotive scenarios. The viewed spectra in each of the lighting scenarios are calculated as follows:

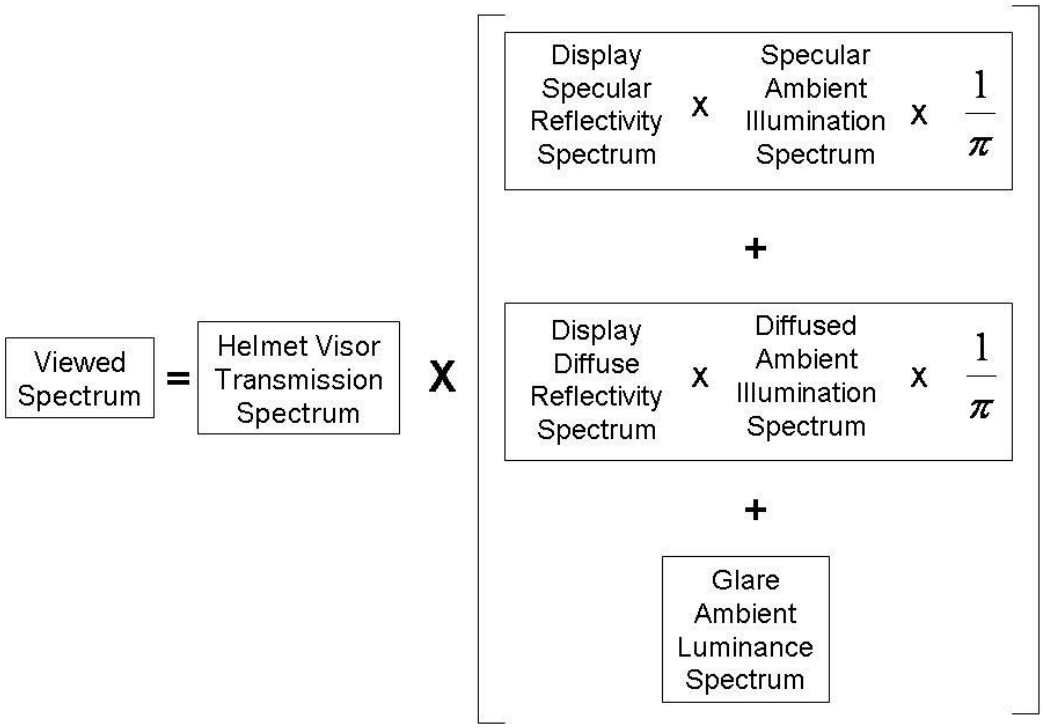

Figure 12 : Calculation of spectral arrays at the eye datum point (BAE Systems 2001)

Figure 12 is a good example of practitioners using just the specular and lambertian component instead of the full BRDF of the display. The charm of this simplified approach is to avoid expensive, complicated and time consuming measurement set-ups and extensive lighting simulations.

By calculating the foreground and background spectra reaching the eye the physical stimuli exciting the human visual system are defined. The next step is to derive the relationship between these physical stimuli and the human perception. In psycho-physical experiments the thresholds for a 50% detection probability of a difference in luminance and chrominance were derived. The threshold experiments were performed by pilots with a line test target subtending 4 arc

minutes on a background tasks luminance of or normalised to 10,000 cd/m². A distance in the CIE 1976 L'u'v' UCS colour space of 0,0042 was found to be the chrominance threshold or one CJND. A common logarithm of the luminance contrast ratio of 0.0051 was found to be the luminance threshold or one LJND. This leads to the following equations for a display's LJND and PJND numbers:

$$LJND_{NO} = \frac{\log_{10}\left(L_{Foreground} / L_{Background}\right) \times RCS}{0.0051}$$

Equation 12: LJND Formula (BAE Systems 2001)

$$CJND_{NO} = \frac{\sqrt{(u'F - u'B)^2 + (v'F - v'B)^2} \times RCS}{0.0042}$$

Equation 13: CJND Formula (BAE Systems 2001)

Where u'F is the foreground u' co-coordinate, v'F is the foreground v' co-ordinate, u'B is the background u' co-ordinate and v'B is the background v' coordinate. RCS is the relative contrast sensitivity as defined by the CIE (CIE 1992).

A symbol is visible when the iso-luminance or iso-chrominance visual difference exceeds the threshold, i.e. is at least one CJND or one LJND. Therefore multiples of LJND or CJND levels are used as a measure of visibility. As most displays exhibit a combination of luminance and chrominance contrasts a combination of CJND and LJND values was defined to be the Perceived Just Noticeable Difference (PJND).

$$PJND_{NO} = \sqrt{(LJND_{NO})^2 + (CJND_{NO})^2}$$

Equation 14: PJND Formula (BAE Systems 2001)

It is a simplification of the psycho-physical process to simply assume that human visual characteristics at threshold level remain true at supra-threshold level. Experimentation within the application field is said to yield reasonable results for the PJND postulation (Vassie 1998). Furthermore it is a simplification to assume that the thresholds obtained for one target size at one relatively high background luminance will be applicable to other scenarios. For aircraft cockpits it is quite reasonable to take the fine lines typical for aircraft instruments as worst case test targets. For automotive applications it is a common question whether visibility of a menu text may be increased by using a bigger font. Therefore a model accounting for target size would be beneficial. Within the calculation of the luminance threshold the CIE1971 Relative Contrast Sensitivity function (RCS) (CIE 1992) can be used to account for the varying contrast sensitivity of the human visual system at different luminance levels. Here the background luminance of the display is used to calculate the RCS. It may be argued that the average luminance of the forward scene should be used instead, as this is the luminance the driver or pilot is adapted to. A problem might be that this is much harder to define and already accounted for by the addition of veiling glare to the foreground and background luminance spectra. Another simplification is the assumption that the CIE 1967 u'v' chromaticity chart really fulfils its aim of being perceptually equidistant. Even though it is much closer to this goal than the original CIE 1931 chromaticity chart the same distance between two colour coordinates in different regions of the u'v'

plane still does not represent the same perception of colour difference. Currently CIE technical committee TC 1-36 is working on a new Fundamental Chromaticity Diagram with Physiologically Significant Axes (CIE 2008).

In spite of all the limitations mentioned above the PJND model provides a highly practical figure of merit. The same PJND obtained in all three lighting conditions means the same visibility in these lighting scenarios even though the display's foreground and background emission power spectra leading to this PJND number may differ significantly. For example a red symbol and a blue symbol having the same PJND are equally visible even though their contrast ratios may be quite different. Furthermore the PJND can be used as a figure of merit as it provides a supra threshold scale for the level of visibility.

By definition any PJND number of more than one PJND leads to visible information. However, Vassie increased the usefulness of the PJND model by establishing a PJND scale linking a required number of PJNDs to visual tasks pilots have to perform in an aircraft.

Task	PJND	Description
Informative	40	Fixed format background information assisting controls or more complex presentations
Status	50	On/off information in fixed location
Static complex	60	More complex information containing small alphanumeric characters and/or fine lines or symbols in fixed locations
Dynamic complex	70	Complex information containing small alphanumeric characters and/or symbols in varying locations
Warning	90	Warning and caution information requiring special attention
Attention getter	120	High priority information which has to be perceived even in peripheral vision

Table 7: Minimum number of PJND levels required for specified tasks in avionics (BAE Systems 2001)

These experiments were performed in an ambient lighting facility (BAE Systems 2001). Pilots adjusted the PJND of various displays within a cockpit mock up to the level they found appropriate for the task to be performed with the instrument.

Car drivers are confronted with information of varying complexity and priority as well. The normal population is less trained on visual tasks than pilots are. Therefore the required number of PJND steps may be different on an automotive PJND scale, but the general approach should be applicable.

3.1.2 Visibility Level

Adrian (1989) used a so called visibility level (VL) to quantify the visibility or non-visibility of targets. Like the PJND number the VL gives the ratio of actual luminance contrast ΔL to the required luminance threshold ΔL_T for just being able to see a symbol.

$$VL = \Delta L / \Delta L_T$$

Equation 15: Visibility Level Formula (Adrian 1989)

< 1 test target not visible

VL = 1 threshold reached

> 1 test target visible

The model is interesting because Adrian uses a more general description for the luminance threshold. Even though Adrian does not include colour contrasts he

accounts for several factors which are ignored by the PJND model. In Adrian's model the luminance threshold is derived from the following parameters:

- adaptation luminance
- target size
- contrast polarity
- duration of presentation
- age of the observer
- veiling glare

The equation for calculation of the luminance threshold ΔL_T is as follows:

$$\Delta L_T = 2{,}6 \times \underbrace{\left(\frac{\varphi^{0,5}}{\alpha} + L_f{}^{0,5} \right)^2}_{\text{Test target size}} \times (F_{CP}) \times \underbrace{\left(\frac{\alpha(L_\alpha, \alpha) + \delta t}{\delta t} \right)}_{\text{Duration of presentation}} \times AF$$

Equation 16: Luminance Threshold L_T given by a scaling factor and the expressions for test target size, contrast polarity F_{CP}, duration of presentation and age factor AF (Adrian 1989)

Test target size

The physical size of the test target is measured in arc minutes and given by the factor α. The other two factors φ and L_f are functions of the adaptation luminance L_a. Here the adaptation luminance is assumed to be the average luminance of the forward scene. For small test targets the threshold becomes smaller with growing

target size. That means bigger test targets are easier to be seen and the target size goes in as a square. However, as targets become bigger the first term reaches zero which means that target size does no longer has any impact on the luminance threshold and adaptation luminance dominates the threshold value. In other words the term for the test target size can be interpreted as follows: For small test targets Riccow's Law [1] holds while for larger test targets Weber's Law[2] takes over and the global luminance threshold is only dependent on adaptation luminance.

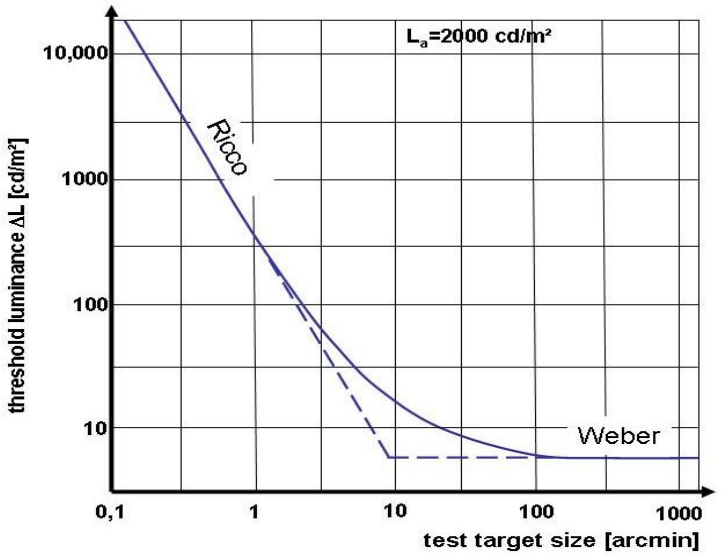

Figure 13: Variance of luminance threshold with target size (Adrian 1989 in Dreyer 2007)

[1] Riccow's Law: Stimuli will be equally detectable if the product of their intensity and area is constant

[2] Weber's Law: The ratio of an increment threshold to the background intensity is a constant

Contrast polarity

Positive contrast is defined as light signs on dark ground and negative contrast as dark signs on light ground. The contrast polarity factor F_{CP} is defined as the ratio of the threshold luminance for negative and positive contrasts.

$$F_{CP} = \Delta L_N / \Delta L_P$$

Equation 17: Contrast Polarity (Adrian 1989)

For the same background luminance test targets with negative contrast are easier to perceive i.e. the threshold for negative contrast is lower than for positive contrast. This accounts for the fact that black text on white paper is easier to read than white print on black paper i.e. $F_{CP} < 1$.

The smaller F_{CP} is the smaller is the global luminance threshold ΔL_T. In detail F_{CP} is a function of the adaptation luminance L_a and target size α. Adrian showed that for very small test target sizes there is little difference in sensitivity to negative or positive contrasts but the difference continually increases up to a maximum of around 25 arcmin and decreases afterwards.

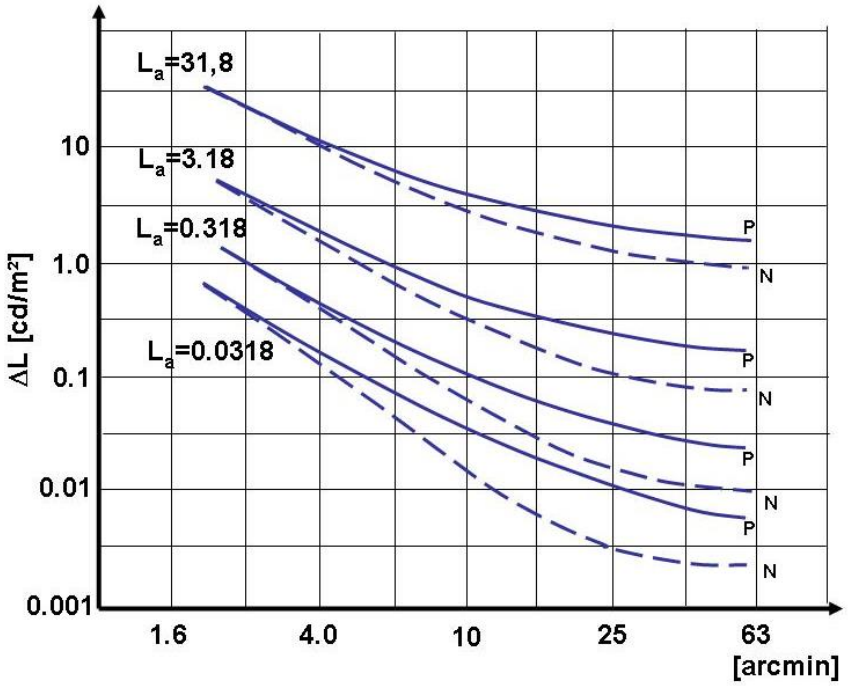

Figure 14: **Threshold for negative and positive contrast polarity (Adrian 1989 in Dreyer 2007)**

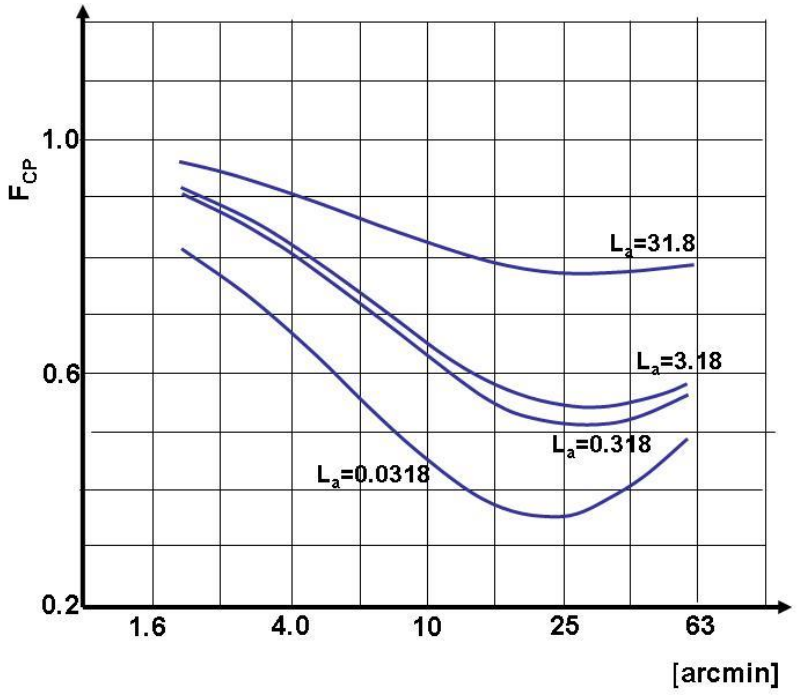

Figure 15: **Sensitivity to contrast polarity of test target sizes (Adrian 1989 in Dreyer 2007)**

This effect is especially pronounced at low adaptation luminance which is quite astonishing as one would expect higher adaptation luminance to have more detrimental influence on positive contrasts (bright writing on dark background) than low adaptation luminance.

Unfortunately Adrian offers no explanation for this phenomenon. As these experiments were performed on purely reflective media like paper it is to be questioned whether this term holds for emissive media as well. Nevertheless it is important to note that this is already the second term in the visibility level equation where target size and adaptation luminance are the main factors and cause significant variance.

Viewing time

The term for the viewing time again depends on adaptation luminance L_a and target size α as well as the viewing time itself. The experiments were performed for presentation times of 2s or longer. For automotive applications 2 s mark the maximum allowed eyes off road time already so shorter times and therefore higher values for the luminance threshold L_T are to be expected.

Age factor

With growing age the eye lenses become more and more opaque and exhibits a scattering effect like frosted glass. Therefore the luminance threshold rises with growing age. This effect is accounted for by the age factor. Figure 16 shows the influence of age on the luminance threshold. The figure exhibits a steep slope just after 60 years.

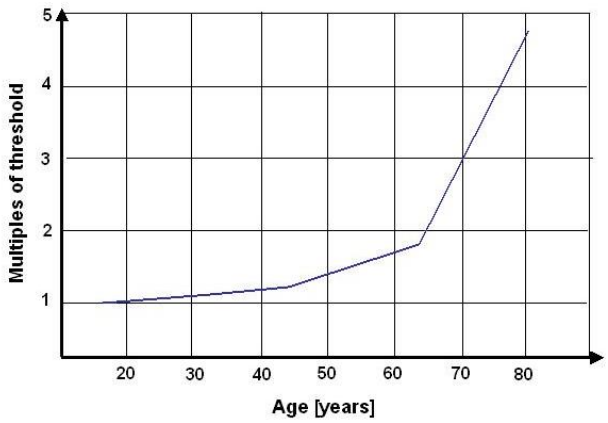

Figure 16: Influence of age on rise of luminance threshold (Adrian 1989 in Dreyer 2007)

The Visibility Level model is highly interesting to automotive applications as it extensively accounts for target size, adaptation luminance and presentation time as well as age factors. Age has extreme influence on the luminance threshold and the age factor is an important difference between the typical pilot and the car driver population. The model can be even further improved by accounting for the difference between global and local luminance thresholds described in CIE Report No. 95 (CIE 1992).

Strictly speaking the contrast threshold calculated is just a global threshold because it is calculated for the whole viewing area. It is based on a homogeneous background with a single target, which is a typical laboratory setup not often encountered in real world scenes. In real world scenes most targets have to be viewed on varying backgrounds and in automotive applications adaptation issues are even more complicated because the driver is adapted to the road-ahead-scenario and then has to look down at his display where the luminance distribution

is likely to be quite different. Dreyer (2007) proposes the following equation to account for the contrast in the direct environment of a test target:

$$C_{Tlocal} = \frac{1}{4} \cdot C_{Tglobal} \cdot \frac{\left(1 + \left(L_u / L_a\right)\right)^2}{L_u / L_a}$$

Equation 18: Contrast in the direct environment of the test target (Dreyer 2007)

Where $C_{Tlocal} = C_T(Lu)$ is the local threshold, $C_{Tglobal} = C_T(L_a)$ is the global threshold , L_u the luminance of the target environment and L_a the adaptation luminance.

The Visibility Level model is highly interesting to automotive applications as it accounts for target size, adaptation luminance and presentation time as well as age factors. When incorporating Dreyer's approach to global and local thresholds, even the differences in the drivers' adaptation luminance and the luminance level in the display area inside the car can be accounted for.

3.1.3 Time to Visibility (TTV) Model

Another visibility model highly interesting to automotive applications is the time to visibility (TTV) model by Silverstein (1996). The model explicitly addresses the impact of dynamically varying viewing conditions on display requirements. This especially includes scenarios where the viewer is adapted to a bright scene and then looks at a display within a darker environment. The corresponding automotive example would be driving in broad daylight with the sun in front and then looking down at the displays which are in the shadow of the dashboard. The driver's eye-sensitivity will adapt from the high outside luminance to the lower interior

luminance, but this so-called transient visual adaptation takes time. Especially in automotive applications, where eyes-off-road time increases the risk of accidents not only the information but whether information is visible/legible at all is of importance. This model offers the benefit of predicting how quickly information can be gathered after a change in the point of visual regard. Therefore the time-to-visibility is a highly valuable figure of merit.

In order to predict the time-to-visibility in dynamic lighting conditions a large number of factors have to be known. First of all the so called intrinsic visual parameters of the display like emitted luminance, contrast, contrast polarity, on/off reflectance and angular display size as well as the spatial frequencies of interest in the displayed information, the level and spectrum of incident ambient illumination and the adapting luminance of fixated regions in the visual task environment. The basic formula for calculating the TTV is as follows:

$$TTV = \frac{ab^{A_{MAX}}}{\tanh\left(C_d - C_{th}(f)\right) / c_{AMIN}}$$

Equation 19: TTV Formula (Silverstein 1996)

Amax = maximum adaptation luminance

Amin = minimum adaptation luminance

DL_{MAX} = maximum display luminance (emitted + reflected)

DL_{MIN} = minimum display luminance (emitted + reflected)

C_d = $(DL_{MAX}-DL_{MIN})/(DL_{MAX}+DL_{MIN})$

$C_{th}(f)$ = contrast threshold at spatial frequency f

a, b, c = estimated model constants

Like in the PJND model DL_{MAX} and DL_{Min} are calculated from the display's emission and reflection spectra and the spectral luminance of incident light. The visual contrast threshold $C_{th}(f)$ is generated by a threshold function after Barten (1992) which scales with the average display luminance and the angular size of the display diagonal. The TTV can be described as the time (in seconds) required for a viewer to resolve a specified level of spatial detail (i.e. spatial frequency in cycles-per-degree) on a display after having been adapted to the luminance of a specified field of foreward view (FFOV. The following table (Silverstein 1996) shows sets of nominal values for incident ambient illumination and FFOV luminance based upon user surveys and photometric measurements in a variety of lighting environments:

Condition	Illumination [lux]	FFOV [cd/m²]
Office (Indoor)	500	100
Conference room	10	171 (bright uv-gap)
Direct Sun Light (outdoor)	100,000	34,260
Car – night	1	0.3
Car – day /lo-ambient	3,000	3,000 (clear blue
Car – day /med-ambient	20,000	10,000 (haze)
Car – day /high-ambient	60,000	34,260 (sunlit
Car – day /hi - FFOV	3,000	34,260 (sunlit

Table 8: TTV Nominal Illumination and FFOV Conditions (Silverstein 1996)

Like the PJND model the TTV model aims to predict readability in certain viewing scenarios.

The contrast threshold for a given spatial frequency of a target in the TTV model is a function of the display luminance and the angular size of the display diagonal. This means it is a local threshold in the sense of the VL model.

Unlike the PJND and VL models Silverstein does not use simple bright/dark contrast ratio but the modulation contrast. Furthermore he brings in the adaptation

luminance which he calls FFOV separately. The PJND model's scenarios are quite similar in the separation of illumination of the display and forward scene luminance. In contrast to the PJND Model, where a veiling glare spectrum is just added to the displays foreground and background spectra, the TTV model really accounts for the dynamic adaptation process. This is an advantage over the static VL model as well. Furthermore the TTV can be used for an ergonomic evaluation of MMI concepts as ergonomists are used to rate tasks by the time users need to complete them. Especially in automotive applications eyes-off-road-time has to be kept to a minimum therefore the TTV is a highly useful figure of merit for driving safety of automotive displays. However, the model has the major drawback of being proprietary i.e. the estimated model constants are not publicly available. Furthermore like all other visibility/readability models the ability to predict the quality of a displayed image is limited as this is based on more factors than readability.

3.2 The Image-Quality-Circle – A Psychometric Approach

The Image Quality Circle is a psychophysical approach to measure image quality introduced by Engeldrum (2000). Engeldrum was looking for a systematic approach to the "print-quality-problem" a company manufacturing photocopiers and printers experienced. The goal was to be able to predict –what impact changes in the components or working principle of a printer would have on the quality of the printing result. The problem was that experiments aiming to form a direct relationship between for example the size of the used toner particles and the quality ratings given by a group of individuals could not be generalised. So he had

a closer look at the processing stages between a technology variable like size of toner particles and the customer's quality preference rating on a printed page. The important step Engeldrum makes is to show that human perception can indeed be used as a reliable meter and is essential to model in order to predict a quality preference

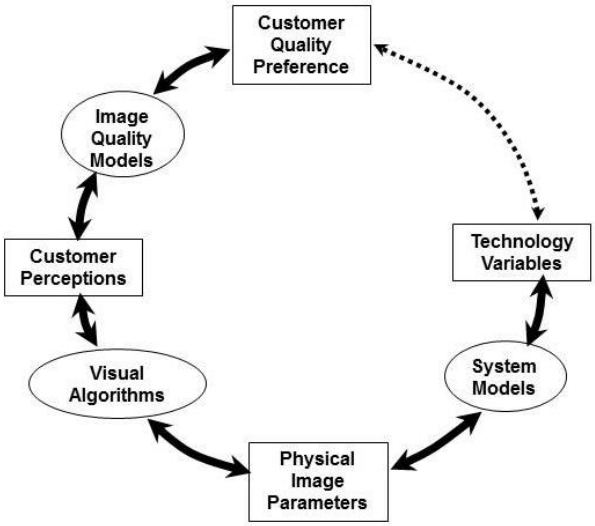

Figure 17: The Image-Quality-Circle (Engeldrum 2000)

One possible starting point is what the customer sees. He does not see the size of the toner particles; he just sees the printed page. The printed page is the physical stimulus reaching the customer's eye. In the Image Quality Circle the physical stimuli are the physical image parameters. A printed page or more generally a physical image can be created by a variety of printer technologies or electronic display technologies or other imaging technologies. Each of these technologies has its own technology variables which lead to a certain form of physical image.

For example technology variables of a liquid crystal display are variables like gap size, tilt and twist angel of the liquid crystals, optical properties of the colour filters and so on. All these technology variables determine what kind of physical image can be produced by the imaging technology, but they are not perceived directly by the observer.

Especially manufacturers of emerging imaging technologies are usually quite reluctant to share the real impact of all relevant technology variables on the physical image. This is understandable, because such companies can't afford to give intellectual property out of the house or even worse to reveal the limitations of their imaging technology. Their own researchers may have system models which can predict the effect of changing a technology variable will have on the physical image produced, but such a model will rarely be shared. However, from the point of view of a company that integrates displays into a product the important thing is the physical image an imaging technology can produce. So typically the physical image parameters to be realised are specified and it is up to the manufacturer to tune the image technology variables in an appropriate way. Another good thing about physical image parameters is that they can be measured by trusted physical instruments in a laboratory. For this reason physical image parameters are often considered as the objective and ultimate description of the properties of an image technology. Basic physical image parameters relevant to automotive displays and their measurement already have been described in chapter 2.2. The only drawback of the wonderfully objective measurement instruments is that they don't spit out a quality rating. To get there human perception has to be brought into the game.

According to the circle of the perception process discussed in chapter 2.1 physical stimuli are only the entrance data to human perception. Researching the relationship between human perception and physical stimuli is the domain of psychophysics. Pioneers of psychophysics like Fechner, Weber and Stevens (Goldstein 2007) have shown that algorithms forming an unambiguous relationship between physical stimuli and human perception of these stimuli exist. As the image quality circle deals with visual perception the link between the visual stimuli or in this case physical image parameters and their perception is called visual algorithms. The contributing aspects to overall image perception are called visual attributes. Since most (not only visual) attributes of perceptions end on ness like brightness, lightness, loudness, sweetness Engledrum also calls his visual attributes lovingly “the nesses”. The visual algorithm linking the visual attribute lightness to the physical stimulus Y for example is the straight forward CIE endorsed lightness function. For other visual attributes like for example sharpness the linking algorithm is less straight forward, as several physical image parameters contribute to the perception of sharpness.

It is according to the general perception process that cognitive parameters like personal preferences and former knowledge of what an image should look like, influence the quality rating. One of the difficulties in establishing the relationships between the individual visual attributes and the overall image quality rating is that this is a many-to-one transformation. All visual attributes present in a perceived image make their contribution to the final quality rating of the image. What complicates the construction of a model is that the individual visual attributes are

not independent. Varying brightness for example has an influence on perceived contrast.

3.2.1 Physical Display Parameters

The input parameters starting the whole perception process are physical stimuli. In this case in first order the physical stimuli given by the automotive display are of interest. These are the **physical display parameters** like luminance, colour coordinates, black level and white level, gamma curve and reflection properties which can be found in display specifications.

Maximum Luminance and Black Level

Luminance is measured with a photometer or a luminance camera. In these instruments a defined display area is imaged onto a calibrated photodetector or CCD sensor. The luminance stated for a display usually is the luminance value for a full white display screen in a dark room, measured perpendicularly to the screen. This is the **maximum luminance** of the display.

Another factor measured in cd/m² is the display's **black level** i.e. the minimum luminance L_{min} which can be achieved when the display is turned on. For most display technologies black level and maximum luminance are not independent parameters. For example increasing backlight power in a transmissive LCD in order to increase maximum luminance of the display increases the luminance of all grey levels i.e. the black level is raised as well. On the other hand, measures to lower a displays black level, like incorporating a dark mask, attenuate the display's maximum luminance. That means for each display there should be a maximum achievable contrast, which is determined by technological display parameters.

Contrast

Basically contrast is a number for the relationship between the brightest and darkest images a display can produce. A commonly found number is **contrast ratio** giving the luminance quotient of full screen white to full screen black measured in a completely dark surrounding.

$$c_r = \frac{L_{\max}}{L_{\min}}$$

Equation 20: Contrast Ratio

Especially for selfluminous displays contrast ratio in the dark is generally the most flattering number. The only number which would be more flattering is the ratio of full white to “display-switched-off”-black in a dark room.
However, real images do have bright and dark parts and the bright parts have the unwelcome nature of lightening up the dark parts. That means the contrast ratio which can be achieved within a natural image on the same display will be lower.

ANSI or “checkerboard” contrast ratio was established to provide a more meaningful contrast ratio figure. ANSI contrast is measured on a test image divided into nine fields of equal size with the middle field white and half the surrounding fields black and half the surrounding fields white. The checkerboard contrast ratio is the average luminance of the white fields divided by the average luminance of the black fields.

The supposedly least flattering contrast definition is modulation contrast. Modulation contrast takes the average luminance of the dark and light contrast pair into account. Modulation contrast c_m is defined as 2 x luminance amplitude divided by average luminance:

$$c_m = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}}$$

Equation 21: Modulation Contrast

Modulation contrast is seldom found in display specifications, but is a very useful number to describe the contrast of elements actually displayed on the display.

Colour Gamut

The area covered when connecting the loci of the CIE colour coordinates in a chromaticity diagram is called **colour gamut**. The colour gamut of a display contains all colours the display can display theoretically. For television the colour gamut -what should be reproduced by all television sets has been normed. One figure of merit for colour displays is which percentage of the NTSC gamut the display covers. The other figure of merit is the bit depth with which the display can be addressed. The bit depth gives the number of colour shades which can be displayed per primary.

Gamma

The **gamma curve** of a display is the function determining how the display's possible dynamic is distributed via the display's grey and colour levels. The name gamma comes from the good old cathode ray tube days where the addressing current was proportional to $Luminance^{Gamma}$. The gamma value linearising the non-linear luminance output from a TV-camera was 2.2. So nowadays monitors are still expected to have a gamma of 2.2 even though for all other displays this response curve has to be simulated. LCDs for example have a completely different original electro-optical response curve but, as a luminance via grey- or colour level curve with a gamma of around 2.2 is the one expected for TV applications and considered appropriate for monitors as well, it is rebuilt by most display technologies' addressing schemes.

However, as the DICOM standard used for optimization of medical displays shows, the gamma curve is not necessarily the best way of distributing the grey levels which can be displayed by a display (NEMA 2003).

3.2.2 Visual Attributes

According to Engeldrum (2000) image quality modelling should be based on the observers' perceptions. In case of displays this are not the physical display parameters but the corresponding visual attributes ("nesses") Research by I. Heynderickx (2005) and Bech (1996) on quality perception of TV sets revealed that the most important visual attributes influencing quality perception were brightness, colour rendering, contrast and sharpness. A short description on the visual attributes lightness is given, because it should be especially relevant to

reflective displays. Gloss is included because it relates to the reflection properties of displays, which have been described extensively in chapter 2.2.2. Homogeneity is mentioned briefly as well, because display specifications often specify luminance tolerances throughout the display in order to guarantee a homogenous appearance throughout the display.

Brightness

The official CIE definition of brightness is: “Attribute of a visual sensation according to which an area appears to emit more or less light.” (CIE 2011)

Thus the visual attribute brightness applies to the perceived luminance of selfluminous displays, which can be thought of as light emitting areas. Furthermore brightness is a so called absolute level of perception i.e. it can be judged on its own and is not related to a surround. Therefore it is also called an unrelated attribute.

Lightness

The official CIE definition of lightness is: “The brightness of an area judged relative to the brightness of a similarly illuminated area that appears to be white or highly transmitting.” (CIE 2011)

In short lightness is a relative brightness. Both CIELAB and CIELUV models use lightness. For reflective media judged next to a perfect white diffuser positioned in a light box giving a known illumination this works quite well. For a reflective display placed in the same environment lightness would be the appropriate visual attribute as well. For a self-luminous display it becomes more difficult to apply the related visual attribute lightness. First of all under typical measurement conditions a

selfluminous display is not illuminated. A common practice is just to assume a D65 white point with Y_n=100. The other possibility is to actually measure the display's white point and use it as reference white. When considering the display in an illuminated environment things become even trickier. Is Y_n the monitor white point or the white of some reflective medium within the scene or a mixture of both? Strictly speaking the adapting white point for comparison of selfluminous and reflective displays under mixed illumination conditions would have to be established experimentally for each new environment.

Perceived Contrast – 'Contrastness'

The Digital Imaging and Communications in Medicine standard defines in part 14 a function describing how the pixel values of a digital medical image should be related to displayed luminance levels (DICOM 2003). This function was named the Grayscale Standard Display Function. The intention behind the development of the Grayscale Standard Display Function was to provide a an objective, quantitative mechanism to ensure visual consistency in the appearance of a given digital image, be it on a computer monitor or on film viewed on a light box.

In order to realise device-independent similarity in the visual appearance of digital greyscale images even between display systems of different luminance the Grayscale Standard Display Function is based on the contrast sensitivity of the human visual system, more specifically the physical model of the contrast sensitivity of the human eye developed by Barten (1992) (See appendix).

$$\frac{1}{M_t} = \frac{1}{k} \frac{\sqrt{T/2}}{\sqrt{\left(1/(\eta p I) + \Phi_0/(1-F(u))^2\right)\left(1/X_0{}^2 + 1/X_e{}^2 + \left(u/N_e\right)^2\right)}} M_{opt}(u)$$

Equation 22: Complete contrast sensitivity function of the Barten Model (Barten 1992)

The fixed constants are: T = 0.1 sec, X_e= 12°, N_e= 15 cycles, Φ_0= 3E-8 sec deg², u_0= 8 cycles/deg, C_{sph}=0.006 arcmin/mm³.

The DICOM standard sets the adaptable values of k, η, σ_0 and p to the following values: k= 3.3, η=0.025, σ_0=0.0133 deg, p= 357 photons/td sec arcmin² which is the p value for illuminant A and uses C_{sph}=0.0001 deg/mm³.

When inserting all the constant values and the concrete values of the DICOM Standard Test Target the complete contrast sensitivity function of the Barten Model reduces to:

$$S(L) = \frac{q_1 M_{opt}(L)}{\sqrt{\frac{q_2}{d^2 L} + q_3}}$$

Equation 23: Contrast Sensitivity Function for DICOM Test Target (DICOM 2003)

with q_1=0.1183034375, q_2=3.962774805E-5 and q_3=1.356243499E-7.

The DICOM Standard Target was chosen to be a 4 cycles/degree sinusoidal grating; because the authors of the DICOM standard found that perceptual linearization for spatial frequencies and object size near the peak of human

contrast sensitivity seemed to work reasonably well for complex images, too. The grating covers a 2 degree x 2 degree area placed in a uniform background of a luminance equal to the main luminance of the target. When viewed from 250 mm distance the Standard Target has a size of about 8.7 x 8.7 mm and the spatial frequency of the grid equals about 0.92 line pairs/mm.

The contrast sensitivity S(L) is defined by the threshold modulation (M_t) at which the grating becomes just visible to the average human observer. The luminance modulation at this threshold represents the Just-Noticeable-Difference (JND) of the Target at the (mean) luminance L.

The Grayscale Standard Display Function is finally obtained by computing the threshold modulation S_j as a function of mean grating luminance and then stacking these values on top of each other. The mean luminance of the next higher level is calculated by adding the peak-to-peak modulation to the mean luminance of the previous level:

$$L_{j+1} = L_j \frac{1+S_j}{1-S_j}$$

Equation 24: Generation of just-noticeable steps in luminance (DICOM 2003)

Def.: One JND is defined as the change in luminance necessary to just detect a difference.

The Grayscale Standard Display Function was modelled for a large luminance range starting as low as 0.05 cd/m² which is the lowest practically useful luminance of CRT monitors and going up to 4000 cd/m² which exceeds the unattenuated luminance of very bright light-boxes used for interpreting X-Ray mammography. The effects of the diffused ambient luminance on the characteristic curve of the display system are explicitly included. Within the luminance range fall 1023 JNDs. The function allows to calculate luminance L as a function of the Just-Noticeable Difference (JND) index, j:

$$\log_{10} L(j) = \frac{a + c \cdot \ln(j) + e \cdot (\ln(j))^2 + g \cdot (\ln(j))^3 + m \cdot (\ln(j))^4}{1 + b \cdot \ln(j) + d \cdot (\ln(j))^2 + f \cdot (\ln(j))^3 + h \cdot (\ln(j))^4 + k \cdot (\ln(j))^5}$$

Equation 25: The DICOM Grayscale Standard Display Function

The index 1 to 1023 of the Luminance levels L_j of the JNDs is given by j. The numerical values of the used constants are: a = -1,2011877, b = -2,5840191E-2, c = 8,0242636E-2, d=-1,0320229E-1, e = 1,3646699E-1, f = 2,8745620E-2, g = -2,5468404E-2, h = -3,1978977E-3, k = 1,2992634E-4, m = 1,3635334E-3

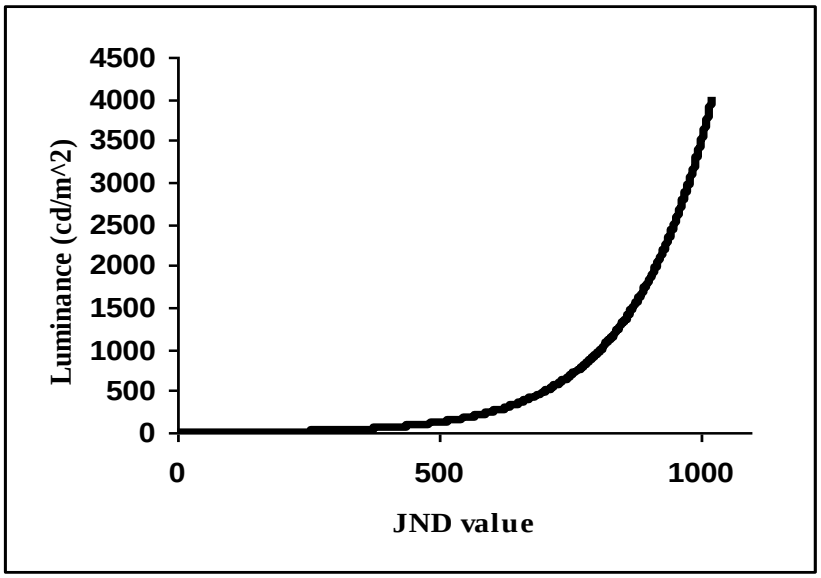

Figure 18: The DICOM Grayscale Standard Display Function (Sharpe 2005)

When characterising a display device with a specific range of inherent luminance values the inverse of the L(j) formula is to be applied.

$$j(L) = A + B \cdot \log_{10}(L) + C \cdot (\log_{10}(L))^2 + D \cdot (\log_{10}(L))^3 + E \cdot (\log_{10}(L))^4 + F \cdot (\log_{10}(L))^5 + G \cdot (\log_{10}(L))^6 + H \cdot (\log_{10}(L))^7 + I \cdot (\log_{10}(L))^8$$

Equation 26: Inverse of the DICOM Standard Display Function

The numerical values of the constants are as follows: A = 71.498068, B = 94.593053, C = 41.912053, D = 9.8247004, E = 0.28175407, F = -1.1878455, G = -0.18014349, H = 0.14710899, I = -0.017046845

Research by Cartwright (2007) uses the concept of linearity in contrast perception to evaluate the sunlight readability of display devices. On the basis of Landolt C's

as test targets observers performed identification tasks of growing complexity. The figure below shows the time after which 95% of the observers correctly identify the test target.

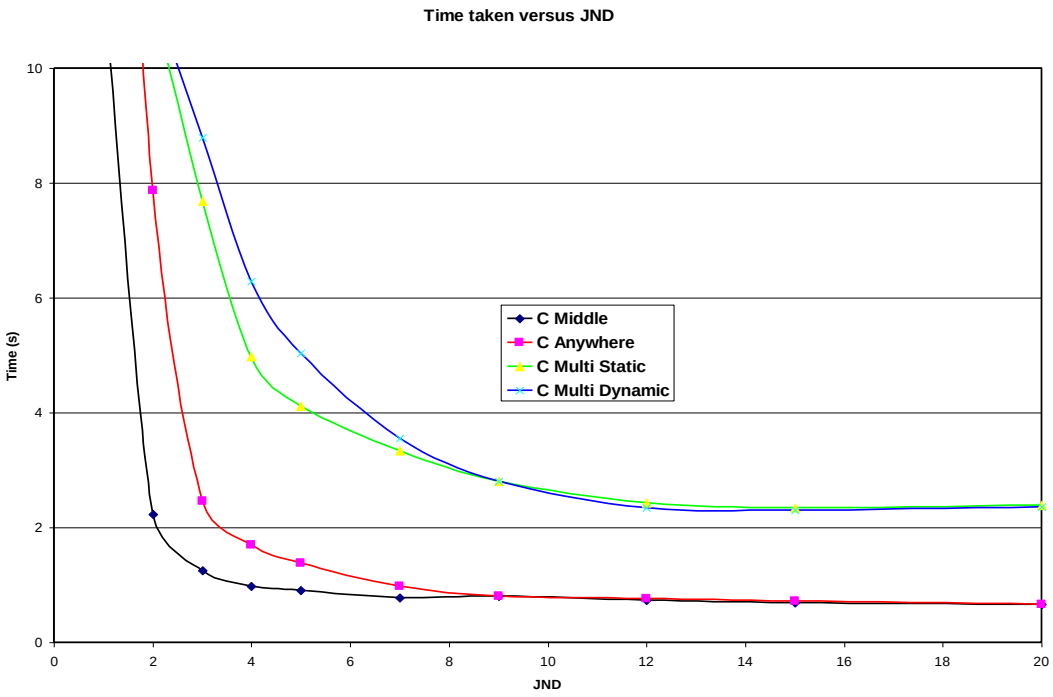

Figure 19: Time taken to correctly identify Landolt C target (Cartwright 2007)

The black line with the blue dots shows the results for simple identification task of static information presented in a known location. In the experiment this is realised by always presenting the Landolt C in the middle of the screen and make observers respond to the orientation of the gap. As figure of merit for the ease of identification reaction time was used. The test series with a middle C showed no further improvement in reaction time for JNDs higher than 5.

The red line with the pink dots represents static information in a formerly not known location, realised by having the Landolt C appear at arbitrary locations

throughout the screen. For this test series identification time settles at the same level as for the fixed location but now 9 JNDs are necessary to reach this reaction time.

The Green line with the yellow marks and the blue line with the turquoise marks both represent information derived by a search task. This is realised by asking the observer to respond to the orientation of the one Landolt C in a number of Landolt C's which has a different orientation. To make the task even more difficult in the blue line series all C's are moving. Interestingly both multi C lines settle on the same reaction time. Due to the added processing step of the search task the reaction time is higher than for the single targets. At least 12 JNDs are necessary to reach the minimum reaction time.

For most applications the minimum reaction time of about 2.2 seconds to respond to information hidden in similar information is no problem. For automotive applications however it is. As explained in the chapter on automotive viewing conditions, the time the driver looks at a display is eyes-off-road time and therefore a safety issue. Glance durations longer than 2 seconds are considered unacceptable and glance durations in the range of 1 to 2 seconds mark a grey region. As the graphic shows a further increase in JND does not improve reaction time, in order to achieve shorter reaction times target size would have to be increased or the complexity of the visual task would have to be reduced.

In short, even for the most complex identification tasks contrast ratios above 12 JNDs do not improve reaction time. Let's apply this result to a colour display built into the secondary viewing area of the driver working place. A transflective display is supposed to deliver 300 cd/m² of white luminance at daylight. That means the background to white writing could be as bright as 277 cd/m² (12 JND steps lower)

and should still ensure readability of the writing. This corresponds to a contrast ratio of only 1:1.083 which is considerably lower than the typically specified minimum contrast ratio of 15:1. Typically contrast ratios specified for other displays or projectors even are in the range of a few hundred. It is assumed that these distinctly supra threshold contrast ratios are motivated by aesthetical considerations. Therefore in the experimental part observer preference on binary automotive symbols with varying supra threshold JND will be tested.

Colour Appearance

A physical parameter commonly used to express the colourfulness of an imaging technology is colour gamut in the CIE1931 or UCS1976 chromaticity diagram. The area covered by the gamut triangle is to represent the colours the technology can produce and is named colour gamut. A typical figure of merit is how much of the area covered by the NTSC gamut the technology can reproduce.

However, to cite Fairchild (2005, p.): "...the display and comparison of the color gamut of imaging devices in chromaticity diagrams is misleading to the point of being almost completely erroneous."

One problem is that colour coordinates provide no information about the colour appearance, because they don't include information about the luminance and chromatic adaptation. On the other hand colour gamut is a widely used figure within the display industry, and can give interesting insights if used properly.

For example psychophysical experiments by Xia et al (2006) investigated to which degree the poor colour gamut of transflective displays can be compensated by higher luminance of the display. They found that generally a higher colour gamut was preferred even at the cost of lower brightness. On the other hand increasing the colour gamut to more than 40% of the ITU REC 709 colour gamut did not bring further improvement. Another result was that image quality rating is affected differently by colour gamut size for highly saturated and barely saturated images.

Sharpness

Bech (2004) and Heynderickx (2005) name sharpness as one of the four most important attributes influencing image quality rating of TV sets. In TV applications sharpness has both a spatial and a temporal component. The spatial component can be understood as how accurate an ideal black/white edge is reproduced. In photography this is typically measured by the modulation transfer function (MTF) of the imaging system which shows how well spatial frequencies in the object are reproduced by the imaging system. The temporal component involves phenomena like motion blur which depends on sampling ratios as well as latencies within the imaging elements.

Homogeneity

Basically all perceptible luminance differences on the whole screen from the real full white image with respect to an ideal completely uniform full white image are calculated and classified in homogeneity evaluation. Homogeneity evaluation is the domain of classical defect detecting models like the spatial standard observer

(SSO) (Watson 2007) or special mura detection and classification models used for automatic quality control of flat panel displays (Chen 2008). Mura is a Japanese term for a blemish and describes a defect in flat panels which looks like a small blurry shape of slightly higher luminance than the surrounding area.

Gloss

The following definition is given by the CIE and takes the dual nature of gloss relating to the perception of images and highlights into account:

“The mode of appearance by which reflected highlights of objects are perceived as superimposed on the surface due to the directionally selective properties of that surface.” (CIE 2011) In more colloquial terms good displays are often spoken of as shiny or brilliant. Both of these terms completely or partially relate to the reflectance properties of the display surface. In chapter 2.2.2 the physical parameters detailing the reflective properties of the display surface are discussed in detail in the section on BRDF and BTDF measurements. Unfortunately it is beyond the scope of this thesis to develop a visual algorithm linking BRDF shapes to perceived glossiness.

3.3 Summary Image Quality Models

In this chapter on image quality models the merits and limitations of readability models in the context of perceived image quality for automotive displays have been discussed at the example of the Perceived Just Noticeable Difference (PJND) model by BAE Systems (2001), the Visiblity –Level (VL) model by Adrian (1989) and the Time to Visiblity (TTV) Model by Silverstein (1996). All three

models give a figure of merit for the readability of the display under ambient illumination conditions. Especially in automotive applications the driver's eyes-off-road-time has to be kept to an absolute minimum for safety reasons. Therefore good and fast readability of the information displayed on the display under all possible ambient lighting conditions is a must for automotive displays. However, once readability is ensured the "beauty contests" starts: A perfectly readable display may still not be perceived as a high quality display.

The Image-Quality-Circle by Engeldrum (2000) described in this chapter visualises the technology independent processing steps from technology variables via physical image parameters to customer perceptions of individual (visual) attributes to an customer quality preference based on the combination of the perceived attributes. Engeldrums systematic approach is used in this thesis and physical image parameters as well as visual attributes relating to display quality were described in this chapter. In the experimental part only brightness and "contrastness" described as JND and ΔJND values after the DICOM metric as well as colourfulness based on gamut size are used. Further investigations on colour appearance in terms of naturalness as well as including the visual attributes sharpness, homogeneity and glossiness would be interesting, but are not part of this thesis.

4 Psychometric Scaling Methods

In the experimental part visual attributes are to be quantified and set in relation to physical display parameters. Hereby the physical display parameters are the stimuli. The perceptions these stimuli evoke in the observer cannot be measured directly by a measurement instrument. For this reason the observer has to be questioned about his perceptions in a reliable and quantifiable way. This is the realm of psychometric scaling and there are numerous different scaling methods available. This chapter gives an introduction to the main types of psychometric scaling methods and some of these will be applied in the experimental part.

4.1 Thresholds and Just Noticeable Differences

In the realm of psychometric scaling there are two types of typical questions to be addressed:

- Can you perceive it?
- Can you perceive the difference?

The first question asks for the absolute threshold of a sensation while the second refers to the just noticeable difference often abbreviated as JND. Practical examples of these in the context of the Image-Quality-Circle are whether for example an unwanted visual attribute like graininess is perceptible at all, or whether a customer will see differences in colourfulness between a reference and a test sample.

In quality and attribute scaling the focus is on obtaining at least interval or ideally ratio scales of steps of just noticeable differences along the scaled attribute. A ratio scale necessitates the zero value of the scale to be known. In case the

absolute threshold of the “ness” is known all values below this threshold can be given the value zero as they are not perceptible.

4.1.1 Direct Scaling Methods

The underlying theory to direct scaling is that humans can assign a number to the magnitude of a sensation and the difference in assigned numbers and perceived magnitude of the attribute will be the same.

There are a vast variety of scales either using numbers or verbal descriptions in use. Figure 20 shows some examples of direct scales:

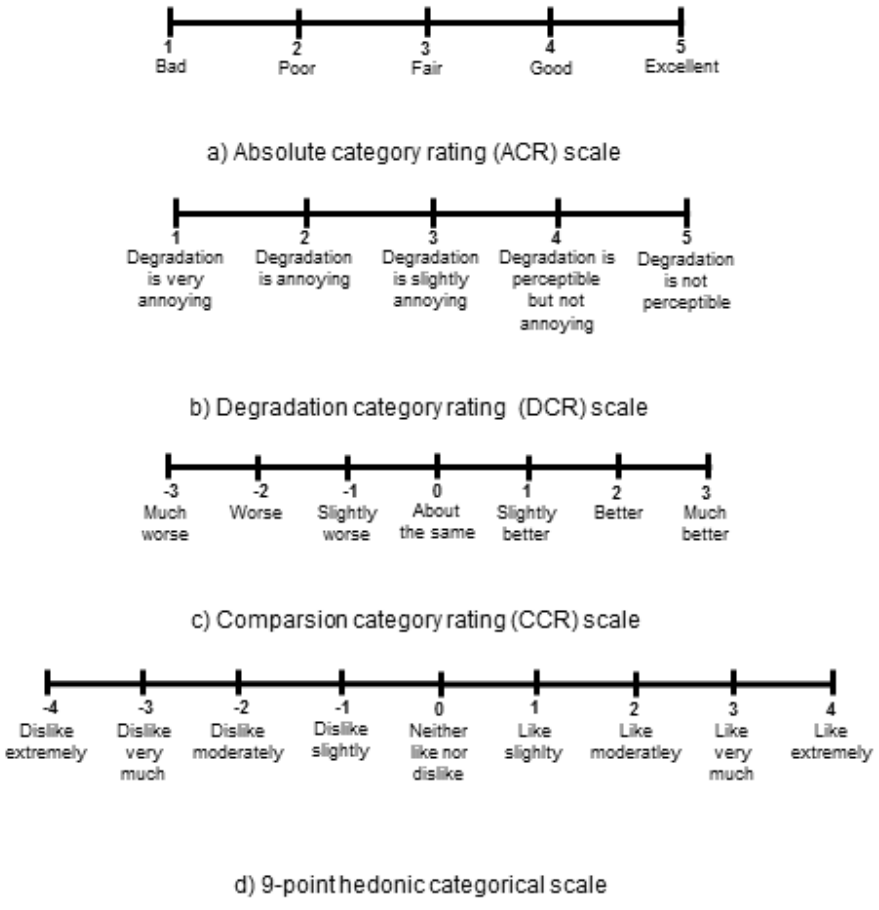

Figure 20: Common Category Rating Scales after Bech and Zacharov (2006)

4.1.2 Indirect Scaling – Paired Comparison Method

In indirect scaling methods the subjects are not trusted to be able to give an adequate number to the perceived magnitude of a sensation. Instead they are only asked to indicate whether samples exhibit more or less of a sensation. A very frequently used indirect scaling method is the paired comparison.

In a paired comparison all possible combinations of stimuli are presented to the observer. The observer's task is to judge which sample of each presented pair possesses more of the "ness" under investigation. The resulting scale is an ordinal scale. For building a quantitative model of the "ness" or image quality it is necessary to convert these ordinal scales into interval scales. This is done by a statistical method called Thurstone's Law of comparative judgement. Thurstone assumes that the difference in choice frequency between samples equals the difference in the magnitude of the perception. Another basic assumption underlying Thurstone's approach is that numbers/choices assigned to a perception follow a Gaussian curve with the real perception being the maximum. The spatial difference between the individual maxima is to represent the perceptual difference. However, this scaling approach only works as long as the Gaussian curves overlap which is only the case when the difference between samples is in the range of few just noticeable differences. As soon as the differences between the samples are way out of the range of just noticeable differences it is not possible to apply Thurstone's Law of comparative Judgement anymore.

4.2 Summary Psychometric Scaling Methods

In this chapter a short introduction to direct and indirect scaling was given. Some common category rating scales were presented as well as the paired comparison method for indirect scaling. In the experimental part the paired comparison method is used for most experiments. Only in the colourfulness scaling and quality rating experiments a degradation category scale is used. The category scaling is used, because too many paired comparison presentations would have been necessary to cover the investigated range of stimuli with close enough spaced stimuli.

5 Reflective Display Technologies

Reflective display technologies are very interesting for automotive applications for two reasons: They excel under high ambient light and quite a few have the ability to show an image in the off state. Especially designers would prefer to show for example a brand logo instead of a black screen in the showroom. The prospect of lower power consumption, improved readability under critical lighting conditions and more design freedom made it interesting to introduce a whole section on reflective display technologies even though only the e-ink display and the high reflective display will come to use in the experimental part.

5.1 Interference Modulated Displays (IMOD)

Micro-electro-mechanical (MEM) devices enable the function of IMODs (Qualcomm 2009). Each subpixel consists of a collapsible reflective membrane and a thin-film stack. The distance between the reflective membrane and the thin-film stack acts like an optical cavity. Ambient light reflected from the top of the thin-film stack will be slightly out of phase from the light reflected off the reflective membrane. The phase difference determines which wavelengths will interfere constructively and which wavelength will interfere destructively. In short via variation of the gap size and composition of the thin-film stack each MEM device can be tuned to the desired colour. The element can be turned black by collapsing the membrane. This is done by applying a voltage to the conductive thin-film stack. When the resulting electrostatic forces cause the membrane to collapse the optical cavity leads to interference in the ultraviolet region. As UV radiation is invisible to the human eye the element appears black.

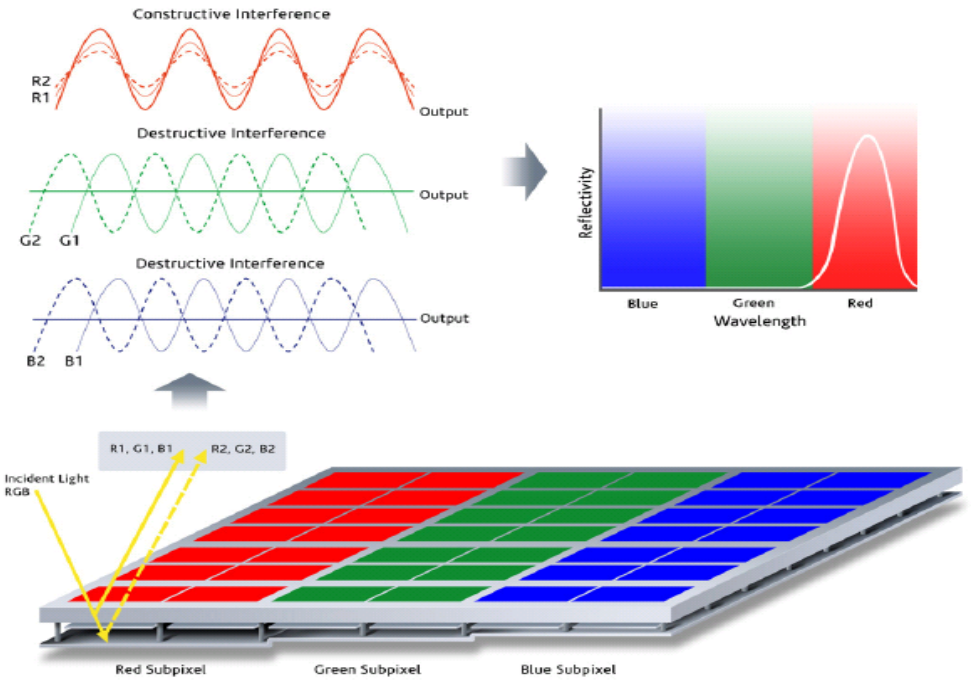

Figure 21: IMOD operating principle (Qualcomm 2009, p5)

The inherent bi-stability of the device enables non-power usage. The bi-stability is caused by the hysteresis behaviour of the electro-mechanical properties. Basically there is an inherent imbalance between the linear restorative forces of the collapsible membrane and the non-linear forces of the applied electrical field. When a constant bias voltage is applied the membrane is held in the open state and the interference colour designed by gap width and the properties of the thin-film stack is displayed. Applying a short positive pulse causes the membrane to be driven into the collapsed state. When the pulse is removed the IMOD element stays in the collapsed state with application of the constant bias voltage. Applying a short negative pulse causes the membrane to snap back into the open state. Theoretically the switching can be done really quickly as gap sizes to be covered

are in the range of nanometres. Considering the difficulties to get hold of an addressable IMOD, driving seems to be not as easy as it sounds.

One problem of the bi-stability arises when not only monochrome, but grey-scale or full colour displays are needed. The IMOD elements do either black or full colour as the inherent nature of bi-stable elements is to have only two stable positions: on and off. This is where the sub-pixels come into the game. In TFT and CRT monitors for example a colour pixel is built out of a red, a green and a blue sub-pixel. While these displays realise shading by addressing the individual sub-pixels with different intensities bi-stable elements have to realise shading via spatial or temporal dithering.

A well-known MEM device using temporal dithering is the digital mirror projector. Here an array of micro mirrors which can be switched to an on and an off state is driven by pulse code modulation. This means within the refresh rate of the image grey scale is realised by flipping the mirror on and off generating as many light pulses as the eye needs to integrate the pulses to the correct continuous grey level. This works extremely well for the digital mirror projectors, but necessitates extremely fast switching speeds and a high mechanical stability of the moving mechanical parts. The drawback of temporal dithering is that the advantage of lower power consumption is lost.

The other possibility to realise greyscale in a bi-stable device is spatial dithering. Like in raster printing with spatial dithering grey scale is realised by the number of sub-pixels switched to the on and off state. The problem with the spatial dithering method is the loss in resolution.

5.2 Electrophoretic Displays (e-ink)

In 2006 one of the few commercially available electronic paper displays was the Sony Portable Reader System PRS-500 (Sony 2006). What made it interesting for experimentation apart from being one of the first products using one of the new reflective display technologies was the fact that it was easily addressable. As an electronic book it was designed for displaying text and had a text magnifying function for sight impaired readers. But what was even more interesting: 4-bit images could be transferred from a computer to the Sony Reader. It accepted the image formats JPEG, GIG, PNG and BMP as well as the common text formats PDF, TXT, RTF and Microsoft Word documents so it was easy to feed the Sony Reader with custom made test images either via the USB 1.1 port or the optical drive for CD-ROM. The screen diagonal was 6” at a resolution of 170 pixels per inch. In short the Sony Reader was a ready-to-go package where the appearance of images and different text sizes could be tested. The only test parameter missing was the timing of the image presentation. A developer’s test kit was available from E Ink Corporation which came complete with a single board computer with Linux and open-source software for display drivers as well as the full source code for operating systems, drivers and applications, offered the timing option, but at ten times the cost of the Sony Reader.

The display technology enabling the Sony Reader is micro-encapsulated electrophoretic ink. The commercial name of the technology is after the company’s name e-ink corporation. The functioning principle is as follows:

Basically charged particles are moved in microcapsules by application of an electric field. There a two possible designs. Either one type of pigment particles is deposited in a coloured fluid or two types of particles with opposite charges are deposited into a clear fluid as shown in the image below. In the first version the bright state is achieved by moving the particles to the front plane where they scatter incident light. The dark state is achieved by moving the particles to the backplane letting the dye absorb the incident light. Greylevels are realised by partial movement of the particles through control of the applied voltage pulse amplitude and duration.

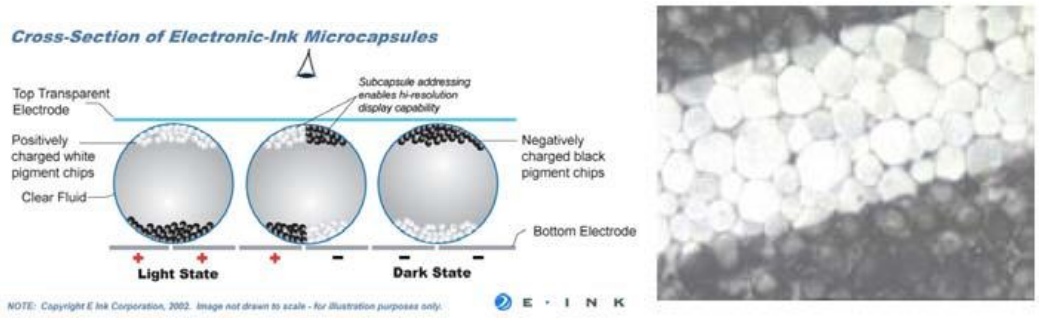

Figure 22: Electrophoretic ink operating principle and microphotograph (E Ink 2002)

The layout used in the Sony Reader uses negatively charged black and positively charged white particles. These particles move in response to an applied voltage across a pixel. Positively charged particles move to the negative electrode and negatively charged particles move to the positive electrode. As the pigment particles are sub-micron size, resolution is in first order determined by the possible resolution of the localized electrical fields forcing the particles to move. In short resolution is down to the quality of the addressing active matrix backplane. Especially where greyscale is concerned which has to be realised by spatial

dithering. The right part of figure 22 shows a microphotograph of micro-encapsulated electrophoretic ink with a mean capsule diameter of 70 µm addressed via a 200ppi active matrix backplane. What can be seen nicely is the sub-capsule addressing.

There are two interesting benefits apart from the paper like appearance: The inherent bi- or multi-stability leading to low power consumption and the possibility to manufacture flexible displays.

Unlike LCDs electronic ink is impulse driven. The optical elements do not respond to the RMS value of an electrical field but to voltage pulses applied with the appropriate amplitude and duration. The polarity of the voltage pulse determines whether the display is driven to higher or lower reflectance values, while amplitude and duration are determined by the difference between the actual and desired optical state. The good thing is that once the desired optical state is achieved, no further addressing is necessary to maintain the image.

The micro-encapsulation and high resistivity of the electrophoretic display material make flexible display set-ups possible. The electrophoretic material is bendable itself. The use of microcapsules compartmentalises the electrophoretic particles thus preventing lateral drift and agglomeration of the optical material. The use of a flexible binder between the microcapsules allows for flexure without permanent deformation of the capsules themselves. Because of the high resistivity of the display material, the current passed by the movement of the particles is very low. This means that relatively poor conductors with resistivity greater than 10^6 ohm/m² can be used. As a consequence the active matrix backplane of an electrophoretic display does not necessarily have to be made out of the typical indium tin oxide

(ITO). Even though ITO is the transparent conductor of choice for most applications it is costly and tends to crack if the local curvature of the substrate becomes too high. Cheaper and easier to bend alternatives are graphite inks, silver inks or conductive polymers.

5.3 Zenithal Bistable LCD (ZBD)

Zenithal bi-stable displays (Jones 2007) realise bi-stability via surface induced bi-stability. They are built just like conventional twisted nematic LCDs with the one difference that one rubbed polymer is replaced by a grating. Coating the grating with a homeotropic alignment layer prevents the liquid crystal molecules from aligning parallel to the grooves as they normally would. Instead the director is normal to the local surface which induces elastic deformation to the system.

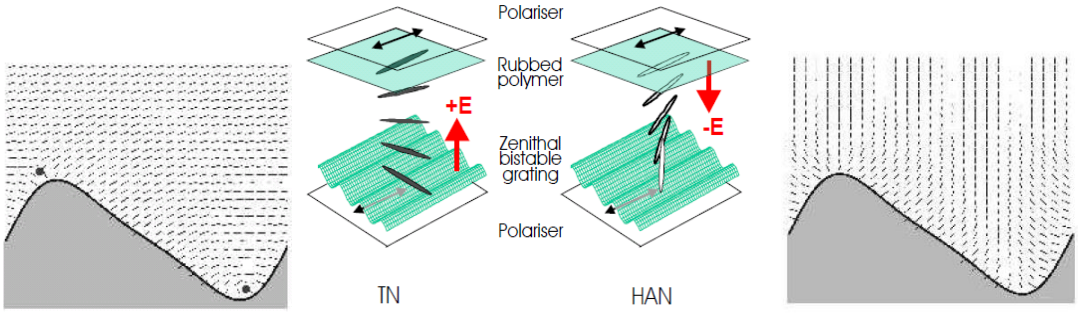

Figure 23: Operating Principle Zenithal-Bistable LCD (Jones 2007)

The degree of deformation can be varied via the ratio of the amplitude a to the pitch L of the grating. The length scales of the grating modulation are on the micron scale in the order to be in the range of the ratio of surface anchoring

energy and the nematic elastic constants. When the gating is shallow i.e. the amplitude is much smaller than the pitch, the molecules in proximity strive to align perpendicular to the grating surface forming a hybrid aligned nematic state. When the grating is deep the hometropic alignment on the steep grooves dominates causing the molecules of the element to align in a planar direction like in the TN mode. This state is supported by a pair of defect lines which occur close to the gratings vertexes. At a ratio of roughly a/L≈0.5 both states have equivalent low energy and the display is bi-stable.

One advantage of the zenithal bistable display is that latching between both stable states appears solely at the grating surface and is completely independent of the orientation of the director on the surface. As a result surface memory effects can't occur making the device free from image sticking. Even more important as the bi-stability is a surface effect it is insensitive to cellgap and temperature variations making the device very robust.

Switching from one state to the other and back is done by applying pulses of opposite polarity. As both stable states exhibit a large distortion of the director field near the grating surface areas of positively or negatively polarised nematic occur. When an electric field is applied normal to the cell a switching torque proportional to the flexo-electric polarisation and the applied field is produced. Depending on the polarity of the pulse this causes defects to be created or annihilated and thus the element to switch into the TN or HAN state. Because of the strong flexo-electric torque the image is retained after removal of the pulse. Latching occurs for pulses of durations of hundreds of microseconds at pulse amplitudes of several volts.

Courtesy of ZBD Displays Ltd one static one bit display sample could be provided. Greyscale is theoretically possible via variation of the grating shape within each pixel but not yet available in a product. The display's viewable area has a size of 82 x72 mm with 320 x 240 pixels at a resolution of 100dpi. That makes a pixel size of 0,234 x 0. 234 mm with a pixel pitch of 0.254 mm which is roughly the dimension of the Sharp HR TFT except for the fact that the pixels in the reflective LCD are subdivided into RGB subpixels. Display update times of typically 750ms at 25°C are still quite long for automotive applications. Furthermore operation temperatures of 0 to +40 °C and storage temperatures of -20 to +70°C are encouraging but would have to be extended for automotive applications as well. The main market for zenithal bi-stable displays is electronic point of purchase (epop) displays. They are used by supermarket chains as electronic pricing and product information which only need power when updated wirelessly.

5.4 High Reflective LCD (Sharp)

Basically the high reflective TFT display is a normal LCD-TFT with a reflecting mirror instead of the back polariser. The RGB filters are mounted directly on the back mirror. It is not bi-stable, but offers 6-bit full colour which is hard to get in a bi-stable display at the moment. The display has 640x240 pixels on an active area of 153,6 x 57,6 mm giving a display diagonal of 6.5 inch. Pixel size is 240x240µm enabling presentation of fine detail. The pixel clock is 14 MHz making image presentation times of about 1 second no problem.

5.5 Summary Reflective Displays

The high reflective LCD display is one of the two display types from this chapter which was actually used in the experimental part. The high reflective display was chosen because, it was fully addressable and had at least 6-bit colour. It was used in the image quality rating experiments as example for the performance of a reflective display technology in comparison to transflective and typical transmissive LCD displays.

The other display from this chapter which was used in the experimental part was the Sony reader, because it was addressable in black and white. This display was used in a matching experiment between reflective and transmissive display appearance which will be described in chapter 7.1.7

iMoD and ZBD and were not used in the experimental part, because freely addressable samples were not available to this project.

6 Daylight Simulation Methods

All readability models described in chapter 3 account for the influence of ambient illumination on display performance. In the real world the source of this ambient illumination is daylight. If the appearance of a display under “real world” conditions is to be reproduced in a reliable way for psychometric scaling of the image quality of the display under “real world” conditions, there are several options available. One possibility is to build a mock-up with high intensity light sources in order to reproduce the relevant lighting scenarios from the overview in chapter 2.2.3 on lighting conditions. Another option is to simulate the daylight appearance electronically. Bases for the simulation of daylight appearance are the CIE sky models described in this chapter. The chapter finishes with a description of daylight mock-ups and facilities which already exist in colour matching, aviation and architecture.

6.1 Calculation of Daylight Appearance

In order to simulate the appearance of a display under all daylight conditions electronically, a mathematical description of daylight throughout the day and in the course of a year is necessary. The CIE provides on the one hand normed spectral distributions of standard daylight illuminants, the CIE illuminants, which make comparability of results easier and on the other hand detailed mathematical models on the composition of daylight, the CIE sky models. Both will be briefly described in this subchapter.

6.1.1 CIE Illuminants

The most important light source is daylight. It can be divided into sunlight and skylight. Sunlight means direct light from the sun casting hard shadows while skylight is the diffuse light scattered by the atmosphere and gives only soft shadows. It is this scattering process which gives the sky its blue appearance. The proportion of sunlight to skylight depends on a number of factors like the nature of the atmosphere and the distance which the light passes through it. The greater the turbidity of the sky for example because of the presence of clouds and the larger the distance the light has to travel the greater the proportion of skylight. Daylight varies severely throughout the day and year due to varying azimuth and elevation of the sun and varying turbidity of the atmosphere due to weather conditions. This results in daylight illuminance on the earth's surface ranging from 150,000 lux on a bright sunny day in summer down to about 1,000 lux on a winter day with a heavily overcast sky. Like the illuminance the correlated colour temperature of daylight varies from 4000 K at an overcast day to 40,000 K for a clear blue sky. Even though models exist which predict daylight incidence on planes at different locations for different atmospheric conditions (Robbins 1986) it is practical to agree on a few standard scenarios. The CIE have agreed on a number of relative spectral irradiance distributions. The most common CIE daylight illuminants are (ISO 11664-2:2007):

- **D65:** Most commonly used daylight illuminant, representing noon daylight at 6504 Kelvin
- **D50:** Warm daylight of 50000 Kelvin, used in the graphics industry due to its even distribution of red green and blue

- **D55:** Mid-morning or mid-afternoon daylight of 5500 K
- **D75:** Overcast daylight of 7500 Kelvin
- **C:** Historical representation of average north sky daylight of 6774 Kelvin

6.1.2 CIE Sky Models

Another phenomenon standardised by the CIE is the distribution of skylight over the sky (ISO 15469:2004). Two main scenarios were defined: a completely overcast sky and a clear sky. For an overcast sky the maximum luminance occurs at the zenith and the luminance distribution is both symmetrical about the zenith and independent of the actual position of the sun. The luminance distribution for a completely overcast sky is given by:

$$L = L_Z(1 + 2\cos\eta)/3$$

Equation 27: Luminance distribution for a completely overcast sky (ISO 15469:2004)

Where L is the luminance of a sky element in L_Z is the luminance of the sky at the Zenith in and η is the angle between the element of sky and the zenith in radians.

The luminance of an element of clear sky can be calculated via the slightly more complicated equation:

$$L = \frac{L_Z (1 - e^{-0.32/\cos\gamma})(0.01 + 10e^{-3\eta} + 0.45\cos^2\eta)}{0.274(0.91 + 10e^{-3z} + 0.45\cos^2 z)}$$

Equation 28: Luminance of a sky element at clear sky (ISO 15469:2004)

Again L_Z is the luminance of the sky at the zenith, but η is the angle between the sky element and the sun, γ is the angle between the sky element and the zenith and *z* the angle between the sun and the zenith.

Depending on the degree of cloud cover, the real luminance distributions of skylight will range somewhere between these two extremes. Average models taken over the complete year were for example established by Littlefair (1985). One can use these values for computer simulation of daylight penetration or try to rebuild these skylight distributions and the path of the sun in an artificial sky. A worst-case scenario sometimes used in architecture is to assume a CIE standard overcast sky producing an illuminance of 5,000lux on the ground.

When aiming to build an artificial sky, it is important to bear in mind that the CIE daylight illuminants are just numbers giving a distribution of relative energy with wavelength representing standard daylight distributions, but they are no real light sources. There are a number of artificial light sources which aim to reproduce the correlated colour temperatures of CIE illuminants these are called equivalent white light sources.

6.2 Daylight Mock-Ups

In this subchapter a few real world mock-ups aiming at reproducing daylight scenarios are described. Light boxes are used for colour matching and are relatively small as only small samples are visually compared inside these boxes. Heliodomes or rectangular skies are bigger and used in architecture to investigate the daylight penetration into architectural models. The biggest sky dome which might accommodate a complete car is the Sky dome by BAE Systems, which is designed to accommodate aircraft cockpits.

6.2.1 Light boxes

Visual evaluation of reflective materials usually is done under controlled lighting conditions. In the automotive industry colour matching of reflective materials is done in lighting boxes like the GretagMacbeth SpectraLight III.

Figure 24: Light box GretagMacbeth Spectralight III (gretagmacbeth 2005)

This box has a dimension of 91 x 61 x 60 (lwh). The walls are made of low gloss neutral grey material to avoid unwanted colour distortions. The lamps, daylight filters and reflectors are placed behind a removable diffuser in the ceiling. As the

lightcabin is designed for colour inspection in a controlled environment ideal for colour vision dimming is not available. For metamerism inspection several colour temperatures can be realized by the use of the following lamps:

Daylight D65 6500 Kelvin:

D65 is an illuminant specified by the CIE as north sky daylight. Here it is realized by a filter tungsten halogen lamp developed by GretagMacbeth. The UV part of D65 is realized by a 6W fluorescent UV lamp. Even though UV does not belong to the visible spectrum, daylight has a non neglectable proportion of UV light which stimulates fluorescent materials to emit in the blue region of the visible spectrum. This method of artificial “whitening” is applied to a lot of materials and substrates for example white paper.

Horizon Light of 2300 Kelvin:

Horizon Light is not a CIE illuminant but a light source specified by the American Society for Testing and Materials (ASTM) in Standard Practice for Visual Evaluation of Colour Differences of Opaque Materials. It is used for metamerism evaluation and meant to represent the red light of the sun at sunrise and sunset. The difference to illuminant A which represents typical light bulbs in domestic environments and has a correlated colour temperature of 2856K is only 556K. For this reason Horizon Light is often replaced by illuminant A. Here horizon light is realized by 4 lamps a 500W to account for the high illuminance values given by direct sunlight.

Iluminant A 2865 Kelvin:

The CIE illuminant for temperature light sources is realized by two light bulbs of 1500 Watt.

Cool White Fluorescent 4150 Kelvin:

Fluorescent light is a common light source in shop or office lighting. Cool white Fluorescent light matches the broadband fluorescent illuminant F9 specified by the CIE. It has a high colour rendering index of 90. The high colour rendering index is achieved by using multiple phosphors which results in a smother spectral power distribution than the rather spiky spectrum of standard fluorescent lamps. The lamps built in are two 30W broadband fluorescent lamps of 4150 Kelvin colour temperature.

6.2.2 Lighting Facilities

In architecture it is common practice to study the daylight penetration into buildings by putting models into artificial sky constructions. These are built to simulate standard overcast sky conditions, giving either uniform luminance or the CIE luminance distribution. The most sophisticated artificial skies use a hemispherical dome structure also called a heliodome.

The left image in Fig. 24 shows the heliodome built at the Welsh School of Architecture (Alexander 2000). The radius is 4m built up by a structure of triangles with 640 luminaries mounted in the middle of each triangle. The lamps used are low energy compact fluorescent lamps with a colour temperature of 4500 Kelvin. Electronically dimmable ballasts provide a dimming range between 3-100%

brightness. At the model stage the dome produces a maximum illuminance of 7000 lux. A slot provides a fixed vertical track with a radius of 4.5 meters for the artificial sun. Three artificial sun types are used at the moment: 1kW tungsten, 575 HMI and a 4kW HMI source all built up as stage lanterns. All these light sources have their individual benefits and disadvantages. The tungsten lamp is easiest to dim, but as a temperature source it has a low colour temperature of 3,200 Kelvin and a cold mirror design is necessary to reduce the infrared heat load. HMI lamps typically provide a colour temperature of 4,500K but can't be dimmed below 30% of their maximum brightness without special equipment. The 4 kW HMI source provides illuminance of over 80,000 lux at stage but makes the use of goggles and high factor sunblock necessary due to its UV radiation. The whole set-up is controlled via a DMW512 system which is a standard in the world of stage lighting. High resolution quadrature encoders are used to determine altitude and azimuth of the track and the object turntable.

Figure 24: Artificial skies: left side heliodome (Alexander 2000), right side rectangular sky (Daiwa House 2006)

A so called rectangular sky is shown in figure 24 on the right hand side. Here only the ceiling is equipped with diffuse illumination. The walls have to be exactly rectangular and equipped with mirrors. The use of exactly opposed mirrors has the benefit of creating the impression of an infinite horizon. The image in figure 25 from the mirror cabinet used at Daiwa House in Japan shows this effect nicely. The room has a dimension of 3.6m x 3.6m x 3.5m. The ceiling is a bit lower due to the space necessary for the lighting devices and the acrylic board used as diffuser. Due to the absorption at the mirror glass occurring at each of the multiple reflections from the diffuse ceiling to the floor a luminance distribution similar to that of the CIE sky is ensured.

An outstanding facility for assessment of display behaviour in daylight is the Ambient Lighting Dome at BAE Systems in Preston, UK. With its 9m dome truncated 2m below the centreline the facility accommodates complete aircraft cockpits. Diffuse illumination levels of up to 25 klux are realised by 72 3kW Xenon arc lights placed evenly all over the outer dome surface. These lamps can be controlled individually and the illuminance is varied by the number of lamps switched on. Direct sunlight is realised via mobile spot lamps of 4 and 2.5 kW. Thus approximately 150 klux of direct illuminance can be achieved without effort. A speciality of this dome is the starlight simulation. This is a purpose-built feature, which has been designed to be used when the aircrew wears night- vision-goggles. The light sources for this special night-time-sky are green and near infrared LEDs providing luminance values of about 0-20 milli lux.

6.2.3 Summary Daylight Simulation Methods

CIE sky models for computer simulation of daylight are available; however including an automotive display scenario in a computer simulation would be a very challenging task. For assessing the visual appearance directly on a real display a physical mock-up seems to be the more appropriate method. While the given examples of lighting facilities stem from other disciplines like colour matching and architecture and are preliminary designed for observers to obtain a visual impression, the displays industries approach is lab based. It is not likely that car makers will invest in a heliodome big enough to accommodate a complete car or even a dashboard mock-up. Typical light boxes on the other hand are a little bit small to accommodate a display within realistic viewing geometries and do not incorporate a very bright directional light source. However, a lab could be equipped to be used as a rectangular sky by building-in a diffuse illuminating ceiling. Current LED technology even makes it possible to tune the colour temperature of such a daylight ceiling seamlessly from 2,700K to 6,500K. Usually a light measurement laboratory features black walls in order to minimise all unwanted stray light which could impair the measurements. However, mounting mirrors to the walls and covering them by a black curtain would give the flexibility to use the lab for dark room and for daylight measurements. Additional light sources for direct illumination according to recommendation DIN/ISO 15008 which is explicitly targeted to measure sunlight readability in automotive applications would complete the “daylight-lab”. This is just a suggestion, a “daylight-lab” has not been realised in this project.

7 Experimental Methods and Results

In chapter 3.2.2 visual attributes were described, which were expected to have an impact on perceived image quality of automotive displays like brightness, perceived contrast (“contrastness”), colourfulness, sharpness, homogeneity and glossiness. The performed experiments concentrate only the influence of the attributes perceived contrast, brightness and colourfulness on observer preference.

The experiments start on observer preference for brightness and perceived contrast combinations for a very simple test target of a bright symbol on a darker background presented on a conventional CRT monitor in a dark room. With a small number of observers (3-5) two short pilot experiments were performed with only 6 image samples, evenly distributed over the dynamic range of the monitor, varying only in perceived contrast or only in in brightness. With a larger sample set chosen with regard to the outcome of the pilot experiments, the main experiments on preference for brightness – contrastness (perceived contrast) combinations were then performed with a higher number of observers (17-20) both in a completely dark environment as well as in an office lighting scenario. The observers in the pilot experiments were members of the display department at BMW group and experienced in image quality assessment. Age of the experienced observers was between 25 and 40 years. The observers in the main brightness - contrastness preference experiments were chosen from engineers of the display group not directly involved with image quality rating, PhD students from other disciplines within BMW Group and a group of technicians performing electronic safety tests in the offices and labs. These observers were technology-

savvy, but not explicitly trained on image quality rating. The age of the observers ranged from 19 to 56 years. The same mixture of observers was used in the colourfulness experiments.

Only one experiment was performed by just one expert observer: In the reflective/transmissive matching experiment an expert observer, trained in photographic colour filtering, tried to adjust the RGB values of two rectangles presented on a PowerPoint slide on a transmissive LCD monitor in order to exactly match the appearance of two rectangles on an illuminated reflective e-ink display. The matching result was then measured by a luminance camera. The idea of this experiment was that, if the appearance of two images was the same, the physical image parameters luminance and colour coordinates should be the same as well, independent of the display technology.

After a discussion of the results of the brightness and perceived contrast experiments colourfulness was the next visual attribute to be investigated. First colourfulness of a transflective automotive display was scaled. To this end saturation of the test samples was varied from 0 to 1 in 0.1 steps and the colour gamut of each image was measured in order to obtain a direct relationship between colour gamut of the display and colourfulness rating. In the following experiments a map image from a driver navigation system exhibiting strong red, green and blue components was presented on the transflective display in varying saturations. The observers were asked to rate image quality of the map image on the transflective display compared to a reference map image on high reflective display and a transmissive display. The reflective display had a significantly smaller colour gamut than the transflective display, but higher reflectivity, the transmissive display had a slightly smaller colour gamut than the transflective

display and brightness comparable to the reflective display. Thus it was possible to investigate to which extent higher colourfulness could compensate for lower brightness or vice versa.

In a first step the influence of perceived contrast and brightness on display quality rating were investigated. This entrance point was chosen, because the DICOM JND formula already provides an algorithm relating white level luminance to a figure corresponding to perceived brightness and the combination of black level and white level to a figure corresponding to perceived contrast. This allows going all the way round the Image-Quality-Circle for the first two parameters and thus tests its applicability to automotive display quality applications.

As derived in detail in chapter 3.2.2 the difference of DICOM JND levels for two luminance values gives the visual attribute of “contrastness” between these luminance levels. This will be called ΔJND. As one JND is defined as a just noticeable rise in luminance in relation to a start luminance, the JND number (short JND) correlates to the visual attribute brightness. DICOM JNDs instead of PJND LJNDs were chosen, because the former are based on the well-recognized Barten model (Barten 1992) of the contrast sensitivity of the human visual system.

These first experiments were done in a dark room to keep range of variable parameters small. One pilot experiment was done in office lighting conditions to show the impact of ambient lighting conditions on quality rating. Geometry issues and timing were excluded in these first experiments as well.

7.1 Supra Threshold Preference Scaling of Contrast Perception

Generally a high contrast ratio is considered a positive figure of merit for display quality and display manufacturers strive to publish the highest possible numbers in their spec sheets. As human contrast sensitivity is distinctly nonlinear JNDs were used for the following preference investigations on contrast ratio. Cartwright (2007) already identified the JND threshold for identification of information to lie between 5 and 12 JNDs (Fig. 18 in chapter 3.2.2), which is considerably lower than typically published contrast ratios of 200:1 or 400:1. Therefore these pilot experiments explored observer preference ratings on distinctly supra threshold contrast ratios.

.

7.1.1 Monitor Characterisation

Due to its image quality and ease of addressing the JND experiments were performed on a CRT monitor (Sony CPD-200SFT, 17 inch, resolution 1280 x 1024). In a dark room the characteristic curve for test targets on a mid-grey (greylevel 128) background was measured in steps of 6 greylevels with a luminance camera. The resulting characteristic curve of the CRT monitor is shown in Fig. 25:

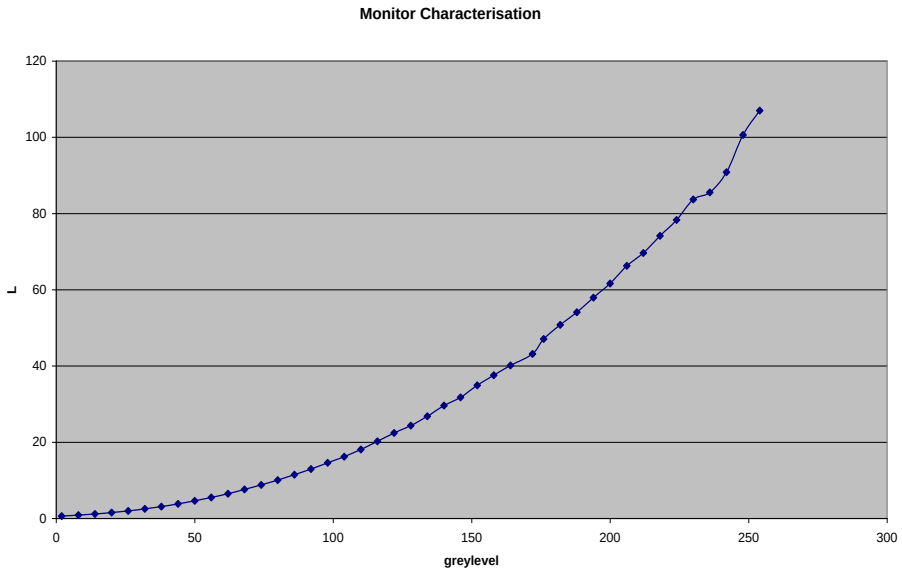

Figure 25: Characteristic Curve of CRT monitor used in JND experiments

The JND values corresponding to the measured luminance values of the monitor were calculated using the inverse of the DICOM Standard Display Function. The result is an already quasi linear JND – greylevel curve for the dark environment. For other ambient illumination conditions a new characteristic curve would have to be measured or modelled and the corresponding JNDs calculated. Basically it would be possible to create internal look-up tables to completely linearise the JND – Greylevel curves for a range of illumination conditions. The drawback is that such a manipulation would minimise the dynamic of the display. At this stage of investigation such an effort was not considered necessary.

7.1.2 Contrastness Preference - Pilot Experiment

The first step was to explore the range of useful stimuli. Therefore a first range of targets with the same mean JND but rising ΔJND were built. The test target was

an automotive symbol in negative presentation i.e. bright writing/symbol on dark background. For all experiments the brighter symbol will be called white level and the background will be called black level, even though for a range of targets one or both elements was distinctly grey. For this first test 6 such test targets were created with a ΔJND between foreground and background of 50, 100, 150, 200, 250 and 300. All test targets had the same mean JND value of 304 JNDs which corresponded to the monitor's mid-grey (gl 128 i.e. 24.37 cd/m²). Three trained observers took part in this first pilot experiment.

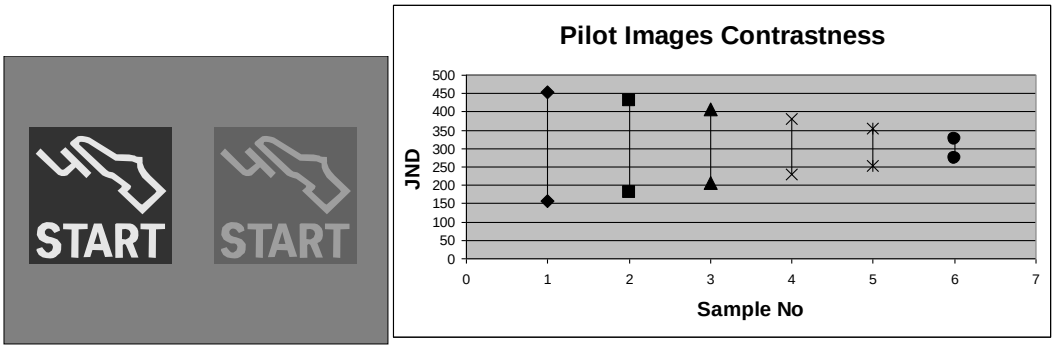

Figure 26: Paired comparison screen and JND distribution of test images

The test targets were presented in the form of a paired comparison as can be seen in figure 26 on the left. For a paired comparison each test target has to be compared with all the other test targets. That means for 6 test targets 15 PowerPoint slides had to be prepared to show all possible combinations once. Because of the small number of first pilot testers and to explore the willingness to assess a larger number of slides 45 slides were prepared which makes 3 runs per observer. The observers were asked to decide for each target pair which one they liked better. The results can be seen in the frequency matrix below:

Sample No	1	2	3	4	5	6
1	0	5	0	3	6	4
2	3	0	3	5	1	2
3	3	3	0	0	3	3
4	3	1	6	0	0	3
5	0	4	3	8	0	0
6	2	4	3	3	6	0

Rank Order	1	2	3	4	5
Sample No	2	1	3	4	5
ΔJND	250	300	200	150	100

Table 9: Frequency matrix and ranking result rising JND pilot experiment

A frequency matrix works as follows: Each time sample i is chosen over sample j a 1 is added in the i'th column and j'th row. Except for sample number one which is the one with the highest ΔJND there is practically no confusion at all between the samples. Observers clearly preferred a higher contrast ratio (ΔJND) over a lower contrast ratio. The only reason for the confusion in the rating of the 300 ΔJND sample was that observers reported the white level to be uncomfortably bright.

Because of the lack of confusion between the samples the results could not be submitted to Thurstone's law of comparative judgement to derive a quantitative preference scale as it is based on quantifying the distances between peaks of Gaussian spreads in judgements. Therefore a simple rank ordering was done as shown in the upper table. In order to get a high enough degree of confusion among the samples JND steps around threshold JND would have to be used. This would lead to an impractically large number of samples to necessary to cover the whole range. The pilot experiment showed a clear trend to preference of higher JND differences as long as the image doesn't get too bright. For this reason a rising ΔJND test with finer steps was not considered necessary at this stage.

7.1.3 Brightness Preference – Pilot Experiment

The pilot experiment on the test targets with the rising JND difference already hinted at an influence of overall JND level on observer preference. To have a first look at that kind of behaviour 6 test targets of the same symbol and a JND difference of 100 JNDs distributed evenly over the monitor's JND range were presented to 5 experienced observers in a paired comparison test. Each sample was identified by a sample number. All possible pairs were presented in random order. This made 15 pairs and this time only a single run was performed. Again observers were asked to indicate which symbol they preferred looking at.

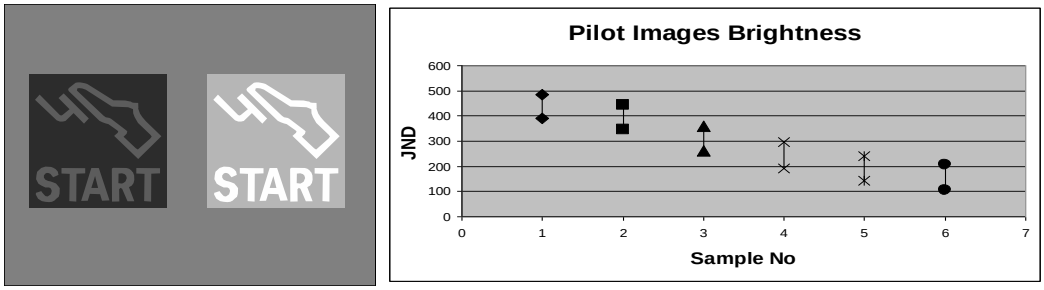

Figure 27: Paired Comparison Screen and Test Target Distribution

Generally observers declared that test targets at high JND levels caused discomfort because of too high brightness and test targets at low JND levels appeared murky. As only one trial with 6 test targets was performed i.e. only 15 pairs rated the experiment protocol is short enough to be show the experimental protocol in table 10:

Sample pair	Subject 1 choice	Subject2 choice	Subject 3 choice	Subject 4 choice	Subject 5 choice
1_2	1	2	2	2	2
2_5	2	5	2	2	2
6_4	6	4	6	4	4
5_1	1	5	5	5	1
3_6	3	3	3	3	3
4_5	5	4	4	4	4
3_4	3	4	4	4	3
5_3	5	3	5	3	3
6_2	2	6	6	2	2
1_6	6	6	6	6	1
4_1	1	4	4	4	4
3_2	2	3	3	3	3
5_6	6	6	5	5	5
1_3	1	3	3	3	1
2_4	2	4	4	4	4

Table 10: Pilot brightness preference paired comparison observer rating

From the paired comparison judgements a frequency matrix of how often each sample is chosen over the other samples is generated like in the first pilot. The frequency matrix over the choices of the 5 observers for the 100 JND Pilot experiment is shown in table 11:

Sample No	1	2	3	4	5	6
1	0	4	3	4	3	3
2	1	0	4	4	4	2
3	2	1	0	3	2	0
4	1	1	2	0	1	2
5	2	1	3	4	0	2
6	2	3	5	3	3	0
∑rel. freq.	1.6	2.0	3.4	3.6	2.6	1.8

Table 11: Pilot brightness preference frequency matrix

The summed up relative frequencies of how often one sample was chosen over the other samples give the rank of the sample. The higher the summed up relative frequencies the better the rank i.e. the better the sample is liked. Table 12 shows the ranking for the set of samples used in the pilot experiment:

Rank Order	6	5	4	3	2	1
□ rel. frequency	1.6	1.8	2	2.6	3.4	3.6
Sample No	1	6	5	2	4	3
100 JND WL	485	206	241	444	296	361
100 JND BL	389	104	143	345	194	262

Table 12: Pilot brightness preference ranking result

In the graphical presentation the summed up relative frequencies of how often the sample was chosen over others are plotted via JNDs for both “black level” and “white level”. The ranking via JND range can be assumed to resemble an upside down U-shape or V-shape with the maximum near mid-grey.

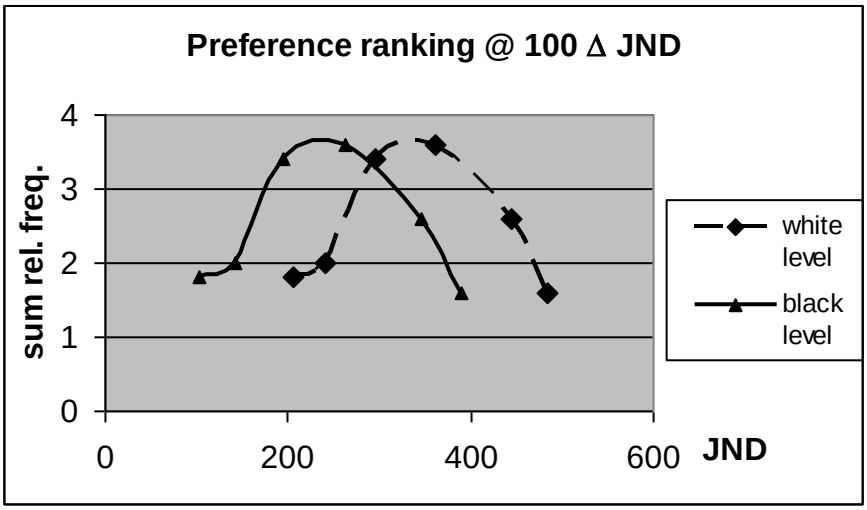

Figure 28: Pilot brightness preference ranking result

The two middle grey test targets were distinctly ranked best. The falling slopes towards low JND levels and high JND levels correspond to observers comments during the experiment. The low JND ranges were reported to be too murky and the high JND ranges to be uncomfortably bright.

7.1.4 Preference for Brightness - Contrastness Combinations

The pilot experiments showed that while contrastness basically follows the rule the more the better, in the dark environment there are limits to the acceptable brightness of the presented symbols. To have a closer look at the trade-offs 12 samples with the same bright automotive symbol (white level) on darker background (black level) covering the whole luminance range of the monitor at three contrastness levels 100, 200 and 300 ΔJND were prepared. A total of 20 observers took part in a paired comparison experiment. Like the pilot experiments the brightness -contrastness preference experiments were performed in a dark room and observers were given time to adapt to the mid grey screen background on which the sample pairs were shown. The mid-grey screen background had a luminance of 24.4 cd/m². The results are given in figure 29:

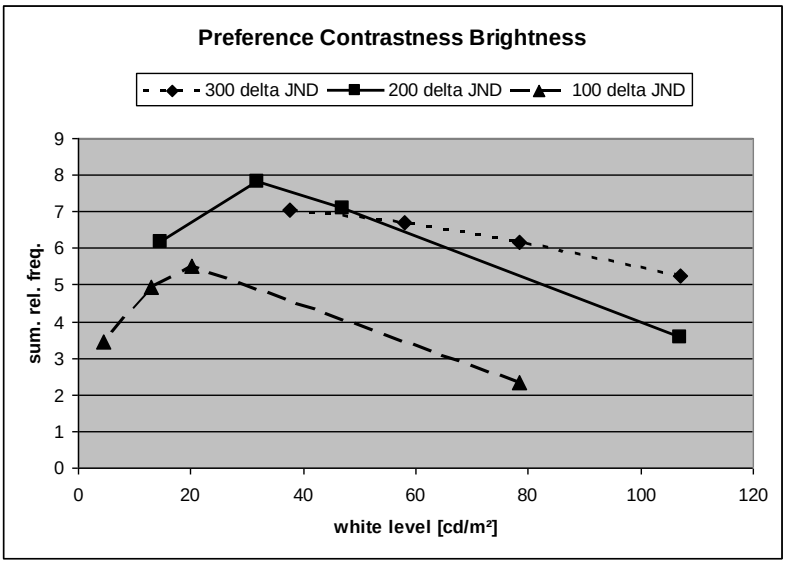

Figure 29: Preference via CRT Contrastness Brightness combinations

The dashed line with the triangles represents the 100 ΔJND samples, the solid line with the squares the 200 ΔJND samples and the pointed line the rhombuses the 300 ΔJND samples.

The preference scaling exhibits a clear trend towards mid-grey presentations. The 200 ΔJND curve shows a comparable behaviour to the 100 ΔJND curve on a distinctly higher preference level. For the 300 ΔJND curve only samples on the falling slope of the curve were producible by the monitor. The slope of the 300 ΔJND curve is less steep and the preference values for the 300 ΔJND curve are not distinctly higher than the preference values for the 100 and 200 ΔJND curve. There seems to be a cut-off point between 20 and 40 cd/m² were higher foreground luminance started to cause discomfort.

The next experiment, performed in the same environment with the same kind of samples was performed to have a closer look at the interesting area around 20

and 40 cd/m². A total of 12 samples were prepared, 6 samples with a contrastness of 150 JNDs and 6 samples with a contrastness of 200 JNDs distributed evenly between a minimum black level of 54 JNDs and a maximum white level of 380 JNDs (corresponding to 47.1 cd/m²). A total of 20 observers took part in the experiment. The ranking result of the paired comparison is given in figure 30:

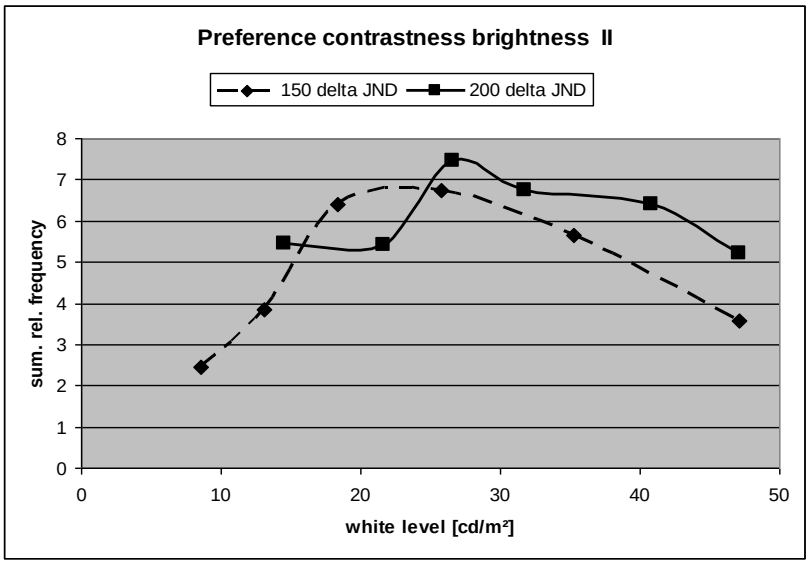

Figure 30: Closer Look at Maximum in Preference Curve

The 150 ΔJND line exhibits behaviour like the 200 and 100 ΔJND curves in the scrambled experiment with a cut-off white level luminance between 20 and 30 cd/m². The 200 ΔJND curve shows an unexpected behaviour at the two dark samples. The darkest sample was rated slightly better than the slightly less dark sample. When consulting the frequency matrix in direct comparison of these two samples the result is the other way round. However, in some of the comparisons the good black at an agreeable ΔJND seemed to overcome the murky impression by the dark white level. There is a little bump in the 200 ΔJND curve created by the

sample with a white level of 31.8 cd/m². This can be explained by the sample having a ΔJND of only 191 JND steps. Even though the lower white level moves the sample nearer to the maximum acceptable white level just below 30cd/m², the decrease in ΔJND seemed to have a greater influence on preference.

As the general behaviour found in the first experiment was confirmed in the finer steps of the second experiment, the next step was to ask observers to rate brightness contrastness combinations relevant to night time driving. The dark room was closest to a night time scenario. The sample types were the same as in the previous experiments, but this time presented on a black screen. The reason for choosing a black screen over a mid-grey screen was to reduce adaptation level. The luminance of full black was 0.4 cd/m² and observers were given 15 minutes to adapt. According to the previous experiments the luminance range of interest was chosen to be up to a maximum white level of 50.2 cd/m². The contrastness range of interest was chosen in 30 ΔJND steps ranging from 200 ΔJND to 290 ΔJND. A total of 17 observers expressed their preference in a paired comparison over a total of 18 samples. The summed up frequencies of how often one sample was preferred over the others plotted via white level luminance is given in figure 31:

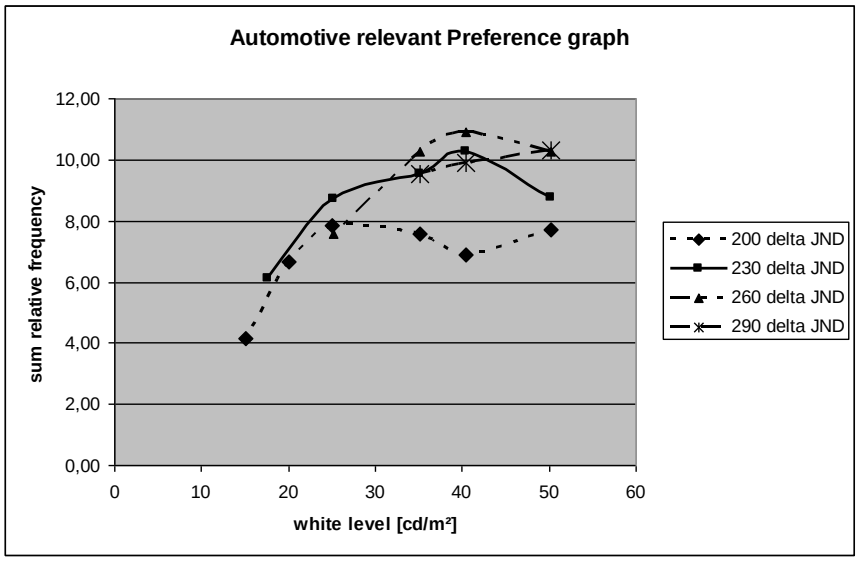

Figure 31: Preference curve via white level of automotive relevant samples

In the white level diagram the curves for 230 (continuous line) and 260 JNDs (dashed with triangles) show a progression similar to those experienced in the previous experiments. Like in the previous experiments the 200 ΔJND curve peaks between 20-30 cd/m² and the higher ΔJND curves peak around 40 cd/m². This means that choosing a black instead of a mid-grey screen background did not have an influence on the maxima of the preference curves.

The pointed 200 ΔJND curve exhibits a new behaviour in this preference graph, a minimum at a white level of 40 cd/m². The 290 ΔJND curve is still on a rising slope while the 300 ΔJND curves in previous experiments already were at a falling slope at these white levels. Both behaviours become less puzzling when plotting preference via black level luminance as given in figure 32:

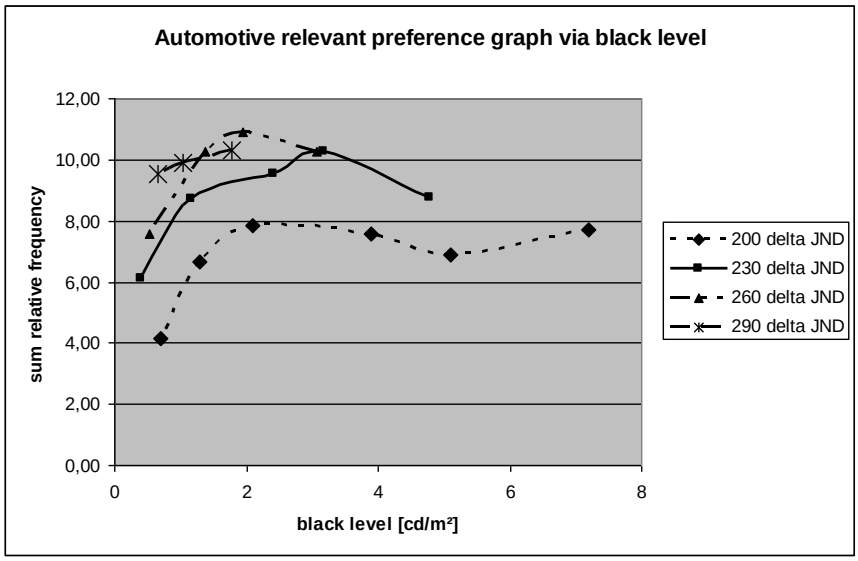

Figure 32: Automotive preference curve via black level

When interpreting figure 32 it is important to bear in mind that rising black levels have corresponding white levels rising by the same amount of JND steps. Generally, except for the 290 ΔJND curve (stars), a higher ΔJND is preferred to a lower ΔJND. As long as the black level is acceptable, preference keeps rising with rising white level. Around a black level between 3 and 4 cd/m² the degradation in blackness begins to exhibit a detrimental influence on preference over all contrastness (ΔJND) levels. The black levels for the 290 ΔJND curve are lower than 3 cd/m². Therefore the 290 ΔJND curve is still on a mounting slope. The pointed 200 ΔJND curve exhibits a minimum at a black level of 5.1 cd/m² and a distinctly higher preference for the next sample with a black level of 7.2 cd/m². A possible explanation for this behaviour is that the black screen on which the samples were presented served as a black anchor. Higher black levels were judged as murky and therefore degraded preference until the samples black level was considerably higher than the black anchor and therefore no more recognized

as degraded black but as an acceptable dark grey. This preference for blackness of black as long as black is recognized as black would also explain the behaviour of the 200 ΔJND curve in figure 31.

7.1.5 Office Environment JND scaling pilot experiment

All previous experiments were performed in a dark room i.e. at a low adaptation luminance. Especially in the experiments covering the whole range of possible luminance and therefore brightness values of the monitor, observers reported to dislike bright samples because they were uncomfortably bright to look at. It was felt that the resulting maximum in the preference curve would be shifted to higher luminance values in brighter environments. Especially the automotive environments can become extremely bright. However, as a first pilot to demonstrate the effect it was judged sufficient to just turn on the normal fluorescent office lamps in the dark painted test room. The resulting illuminance on the screen was 200 lux. The monitor was characterised again under this new illumination condition. To maintain comparability to the experiments in a dark room the ΔJND series were chosen to be 300 ΔJND, 250 ΔJND and 200 ΔJND. Ten samples with the same symbol as in the previous experiments were distributed over the whole luminance range of the monitor.

A total of 6 observers took part in the office lighting experiment. These observers had performed the night time scaling experiment before and performed the office environment experiment after a brief break. The ranking results are given in table 13:

Rank Order	1	2	3	4	5	6	7	7	9	10
Sample No	1	4	2	5	6	3	7	9	10	8
□□rel. freq.	7.8	7.3	6.0	5.8	4.7	4.0	3.0	3.0	2.0	1.3
WL [cd/m²]	102.0	80.3	102.0	80.3	80.3	102.0	60.3	51.8	51.8	60.3
BL [cd/m²]	6.4	4.1	19.3	7.9	14.0	19.3	5.0	3.8	3.8	9.4
□JND	300	300	250	250	200	200	250	250	200	200

Table 13: Rank Order in Office Environment

When plotting the summed up relative frequencies of how often one sample was chosen over the others via white level three nicely stacked preference curves emerge:

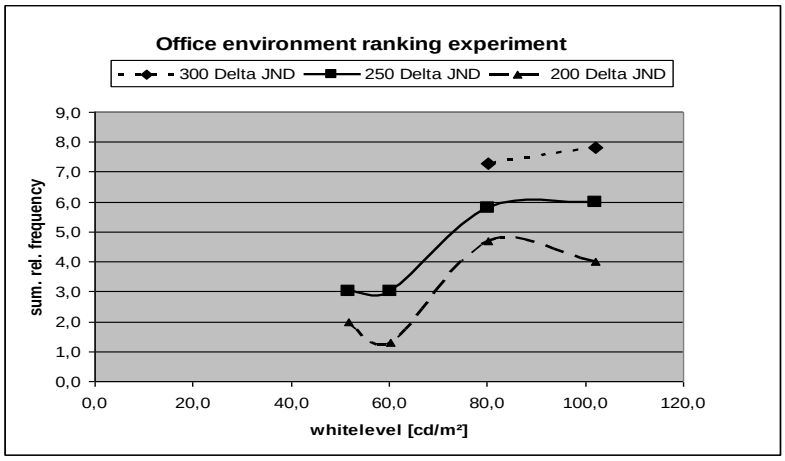

Figure 33: Scaling Result Office Environment Pilot

The first obvious difference to previous experiments is similar distances between the 200 ΔJND, 250 ΔJND and 300 ΔJND curves indicating a linear rise in preference with rising contrastness expressed as ΔJND. This leads to the conclusion that generally there is no detrimental influence of uncomfortably bright white levels in the brighter office lighting set-up. The 200 ΔJND and 250 ΔJND curves both exhibit a minimum and a maximum. However, as expected the maxima are shifted considerably to higher white levels of 80-90 cd/m² compared to

the maxima between 20-40 cd/m² in the dark room experiments. The minima in the 200 ΔJND and 250 ΔJND curves depict the trend that blacker black is preferred to murky black.

The preference ranking experiments for contrastness-brightness pairs showed the same basic behaviour in a dark and a lit environment with a tendency for preference of brighter samples in the lit environment. In all experiments contrastness expressed as ΔJND, brightness expressed as white level JND and blackness of black exhibited an influence on preference ranking. Because of the known relationship between JND level and luminance preference graphs could be plotted against the familiar physical parameter luminance instead of the perceived visual attribute brightness (given as JND level). Thus on the basis of these few visual attributes the link between customer preference and physical image parameters could be formed in agreement with the Image-Quality-Circle process.

7.1.6 Reflective/ Transmissive Matching Experiment

The aim of the reflective/ transmissive matching experiment was to show technology independence of visual appearance i.e. the same physical stimuli produce the same appearance independent of the display technology employed to produce these stimuli.

A Sony reader e-book incorporating an e-ink display as described in chapter 5.2 was used as an example for a reflective display technology. The display was capable of 4 greylevels. Two test targets were built for the e-book: a white rectangle on black background and a light grey rectangle on dark grey

background. The task was to match luminance and colour coordinates of the corresponding fields in the corresponding images on the CRT monitor used in the previous experiments to the appearance of the rectangle images on the reflective display.

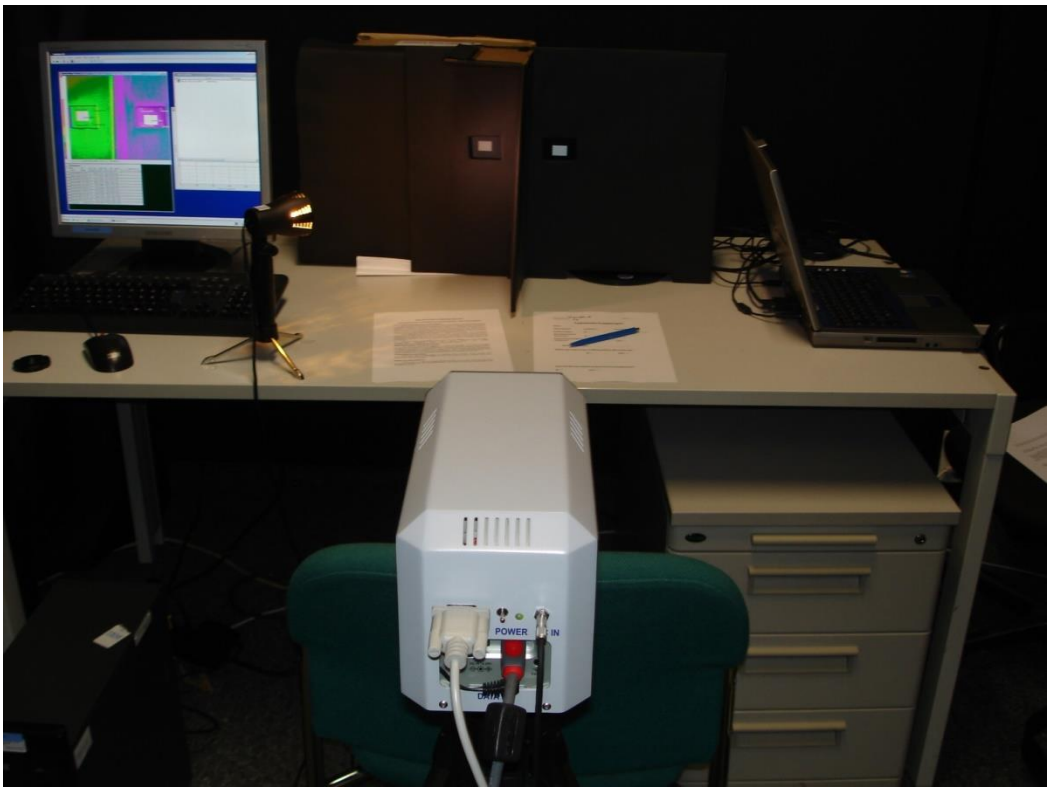

Figure 34: Experimental set-up reflective/transmissive matching experiment

The e-book was placed next to the monitor and both were covered with a blend of black cardboard. The blend had two rectangular holes of the same size one for the reference image on the e-book and one for the image on the monitor. The reflective display was uniformly illuminated with a 500W halogen light source. On the monitor the test image was built as two rectangles on top of each other in a PowerPoint slide.

Figure 35: Image of luminance camera for black white match with and without black blend

An observer experienced in photographic colour filtering adjusted the PowerPoint image palette RGB values for each rectangle until the subjectively best achievable match. This match was measured by a luminance camera. Figure 35 shows the camera image for the black/ white match and table 14 shows the averaged measurement results for the black/white and light grey / dark grey matches.

Spotmeters:	size	Avg	x	y	z	cct
name	[pixel]	[cd/m²]	[1]	[1]	[1]	[K°]
e-book black	20	31	0.4379	0.4066	0.1555	3004
monitor "matched black"	20	27	0.4357	0.4087	0.1556	3058
e-book white	40	115	0.4430	0.4119	0.1451	2963
monitor "matched white"	40	90	0.4226	0.4096	0.1678	3301
black blend	45	28	0.4527	0.4112	0.1362	2808
dark grey e-book	20	74.5	0.4408	0.4126	0.1489	3041
"dark grey" monitor	20	67.1	0.4380	0.4138	0.1483	3061
light grey e-book	40	99.7	0.4422	0.4101	0.1477	2961
"light grey" monitor	40	82.9	0.4390	0.4130	0.1480	3037
black blend	45	26.5	0.4536	0.4092	0.1373	2779

Table 14: Measurement results match monitor to e-book black/ white

Especially for the dark grey /light grey matches the colour match was extremely good. However, the luminance match for the dark grey/ light grey match was less satisfactory. A difference of 17% in light grey luminance and 9 % in the dark grey luminance is considerable. A deviation of 22% for the white match is even worse.

While the match in colour coordinates, the physical parameters corresponding to colourfulness, was quite good, the match in luminance, the physical parameter corresponding to the visual attribute brightness, was not as good as expected. The inaccuracy in brightness matching seems to increase with the brightness of the targets to be matched. For all fields of the same perceived brightness luminance on the monitor was lower than luminance of the corresponding illuminated field on the reflective display. Further experiments, including a closer look at adaptation issues, would have been necessary to reveal if there was a significant difference in brightness perception of reflective and transmissive displays. Exploring this was abandoned in favour of including experiments on the influence of colourfulness on image quality rating.

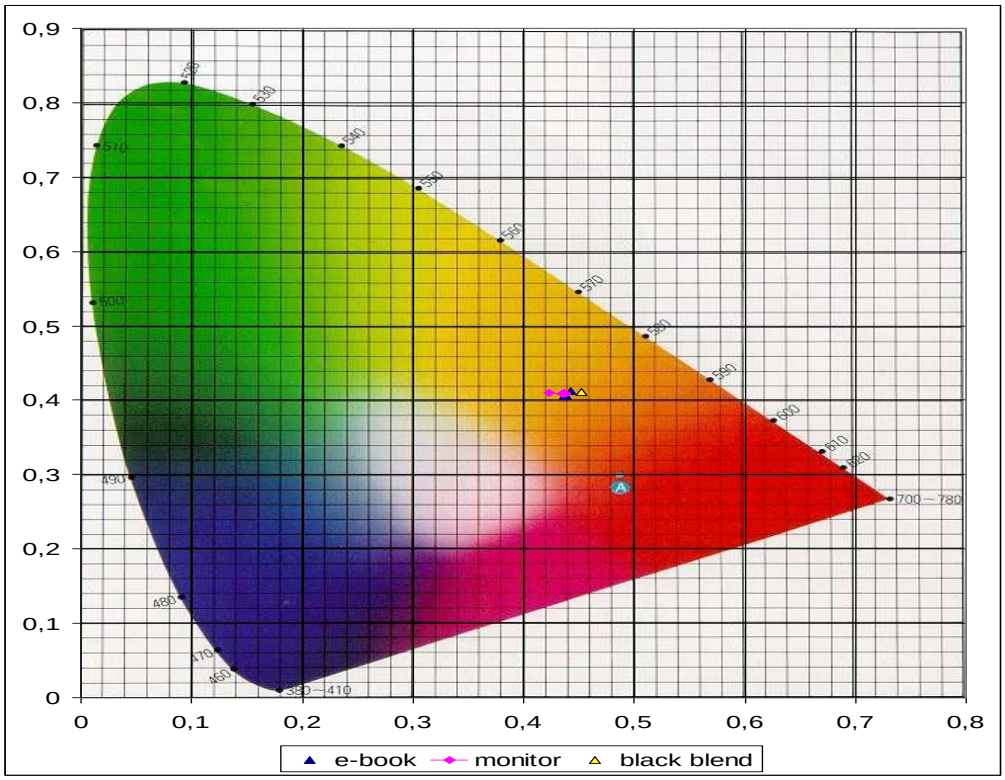

Figure 36: Colour coordinates match monitor to e-book black/ white

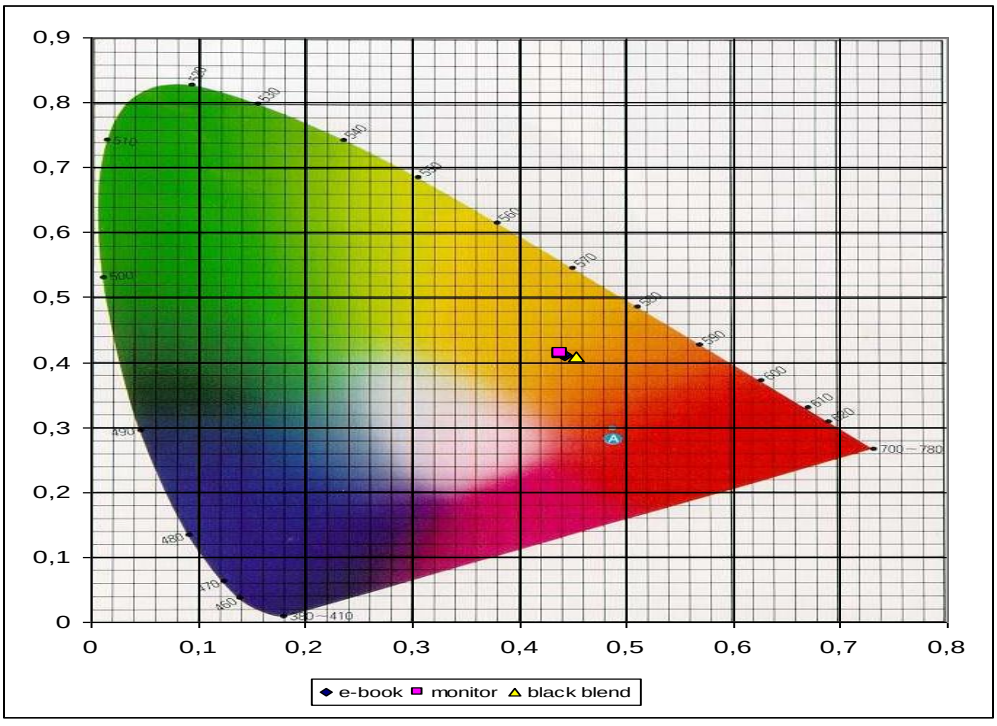

Figure 37: Colour Coordinates matched monitor to e-book greys

7.2 Discussion of JND Results

The main result from the JND scaling experiments was that even though a given JND difference is supposed to deliver the same perceived contrast over a wide luminance range, observer's preference does not stay constant throughout the luminance range. Instead preference rises with rising luminance level until it reaches a "cut-off" point where the bright parts of the image become too bright and preference rating decreases again.

A similar behaviour is reported by for office lighting. In a study by Muck and Bodmann reported in Boyce (2003) a group of 20-30-year-olds were asked to find a specific two-digit number out of 100 similar numbers printed in black ink on gray paper and laid out at random on a table. The search task had to be performed at illuminance levels varying from less than 100 to 10.000 lux and mean detection speed and percentage of observers considering the lighting good were recorded.

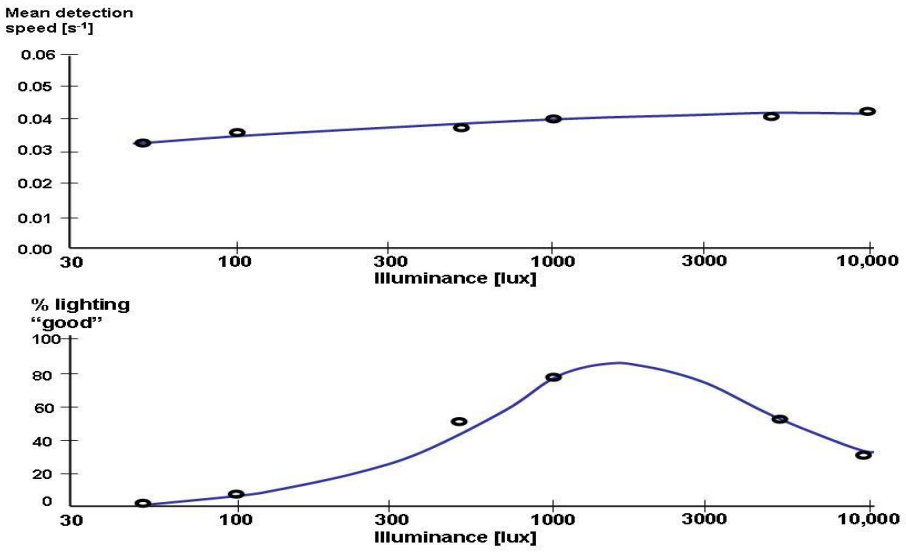

Figure 38: Detection speed versus liking of task illuminance (Boyce 2003, p.187)

Figure 38 depicts the mean detection speeds for locating a specified number amongst others at different illuminances and the percentage of subjects who consider the lighting “good” at each illuminance. Detectability stays practically stable throughout the observed illuminance range (in the Muck and Bodmann experiment it even exhibits a slight continuous rise) while preference rises significantly until an illuminance of about 2000 lx above which observers report the lighting to be uncomfortable and acceptance drops considerably. This behaviour corresponds well to the behaviour found in the brightness preference pilot experiment described in chapter 7.1.3 and the preference for brightness-contrastness combinations experiment described in chapter 7.1.4 In these experiments presented symbols were distinctly supra threshold i.e. detectability was easily given for all presented symbols. Image quality ratings rise with higher contrastness and foreground brightness, but degraded as soon as the bright foreground symbol was perceived as being uncomfortably bright.

Throughout the performed JND difference and range ranking experiments a higher JND difference is generally preferred to a lower JND difference as long as the discomfort region is not reached. Samples with the same JND difference however did not receive the same preference rating throughout the investigated luminance range, but preference exhibited behaviour comparable to the lighting rating experiment with a rise until a saturation or cut-off point above which the viewed image is considered uncomfortably bright.

It is to be suspected that one reason for the decline in preference after a certain maximum luminance is reached is discomfort glare. Typical formulas developed by

the lighting industry for quantifying discomfort glare are not applicable to looking straight at a display because the angle between the glare source and the forward line of sight is placed in the denominator. However, Fig 39 by Hopkinson and Colins 1970 (in Boyce 2003.page 61) classifying shadow, discrimination and glare regions for object luminance by adaptation is much more tolerant about at which object luminance the glare region begins:

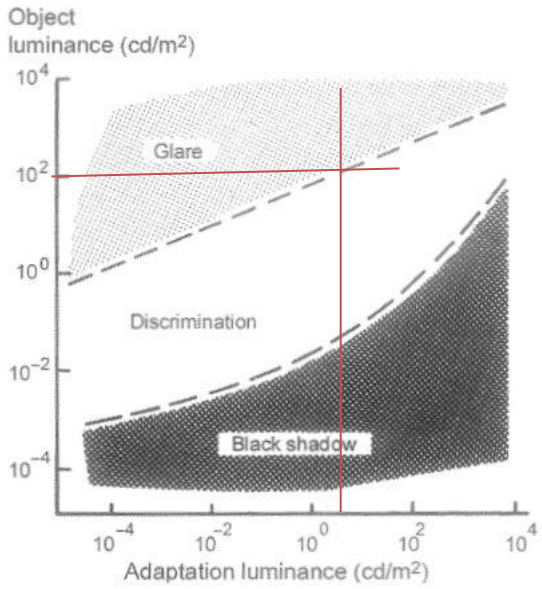

Figure 39: Shadow, discrimination and glare via adaptation luminance (Boyce 2003, p.61)

Figure 39 is a schematic illustration of the range of object luminance within which discrimination is possible for different adaptation luminance. The boundaries are approximate. The red line indicates the adaptation luminance of the observers which were adapted to the mid-grey monitor background of 24.37 cd/m². The glare region, only starting above 100 cd/m², can in the above graphics can be explained by the different question underlying the research. In the experiment by Hopkinson

and Collins (1970) a task performance is the classification criterion. Object luminance is to be expected to become uncomfortable a long way before it becomes so high that discrimination of luminance objects is no longer possible. When assuming that discomfort glare follows the same run as the glare border indicated above it can be expected to rise linearly with adaptation luminance on a double logarithmic scale.

On the other hand the night time scaling experiment clearly shows that discomfort glare is not the only parameter influencing image preference. The number of JNDs covered the black level as well as the perceived distortion of reference values do contribute to preference ranking as well. For this reason the analysis of the data through a multivariate analysis might yield even better results than simply summing up the relative frequencies of choices.

The office environment pilot performed in chapter 7.1.5 showed principally the same behaviour of brightness – contrastness preferences than the experiments performed in a dark environment: Higher perceived contrast is preferred to lower perceived contrast and within perceived contrasts significantly smaller than the dynamic range of the display mid-grey contrast are preferred the same perceived contrasts at the upper or lower end of the display's dynamic range. Distinct minima for background luminance levels which are just slightly higher than minimum display luminance show that the importance of perceived blackness for image quality rating is not limited to dark environments, but is an important attribute influencing image quality in bright environments as well. The main difference to the experiments in the dark environment was that due to higher adaptation luminance,

the preference curves were shifted towards higher luminance values and preference for higher perceived contrasts does not get degraded by discomfort glare, because the maximum luminance of the used display did not reach the discomfort glare region in the office lighting environment. Compared to worst case illumination scenarios in an automotive environment, office lighting conditions still represent a relatively low adaptation luminance and illuminance on the display. For this reason in real automotive daylight scenarios an even more significant shift towards preference of even higher ΔJND and maximum JND values is expected. Even though it could not be investigated in this thesis, it is furthermore expected, that optimum perceived contrast and brightness will strongly depend on the viewing scenario.

As experiments under representative daylight scenarios for automotive displays are out of the scope of this thesis the next experiments focus on the visual attribute colourfulness. Colourfulness has already been named as one of the visual attributes influencing image quality rating in chapter 3.2.2.

7.3 Colourfulness Experiments

A typical figure of merit for how much colour can be produced by a display technology is how much of the NTSC colour gamut is covered by the display technology gamut triangle in a chromaticity chart. Engineers love this practical 2 dimensional representation of an imaging technologies colour gamut even though currently a number of experts start to fight this representation because of inaccuracy (Fairchild 2004b, Brennesholtz 2006, Poynton 2007). The main critique points to current practice are that a two-dimensional presentation lacks to convey

the three-dimensional aspect of colour appearance (i.e. the brightness/lightness influences) and that the NTSC broadcast colour gamut, invented in the year 1953 has never been much in use. The current worldwide recommendation for digital high-definition TV is ITU Recommendation BT.708-2. This is the standard which is nowadays in use and it defines a colour gamut which covers only 71% of the NTSC colour gamut when traced in CIE1931 x, y coordinates. For UCS u', v' coordinates the figure is a bit more favourable; here the Rec. 709 gamut covers 88% of NTSC gamut. As the ideal solution to the named dilemma Poynton suggests to express gamut volume in cubic ΔE_{ab} units.

It has been shown in chapter 2.2.2, that the ΔE_{ab} colour difference formula does not perform as perceptually uniform as intended. Improved colour difference formulae such as ΔE_{94}, CIEDE or ΔE_{99} have been established (CIE 2001, DIN 2001). However, as this thesis is not about uniformity of colour spaces and colour difference formulae, the well-known and widely used CIE1931 xy chromaticity chart will nevertheless be used to describe the input stimuli to the colourfulness experiments.

The goal of the colourfulness scaling experiments was to establish how colourful a display of a given colour gamut is judged. In other words a display colourfulness scale based on the physical display parameter colour gamut was to be established. In the Image-Quality-Circle this would be a visual algorithm linking the physical display parameter colour gamut to the visual attribute colourfulness.

After the colourfulness scale was established the final image quality rating experiments were performed as a comparison between the image quality of a transflective LCD as compared to a reflective and a transmissive LCD. The test image was a representative navigation map image. Brightness and colourfulness of all three display technologies was different and colourfulness of the transflective display was the visual attribute varied in the image quality rating experiments.

7.3.1 Colourfulness Scaling

The first colourfulness scaling experiment was done in a dark room on a common laptop liquid crystal display. The maximum colour gamut of the display relative to the NTSC colour gamut, a common CRT display, a laser projector and the transmissive part of a transflective automotive display is given figure 40 and table 15:

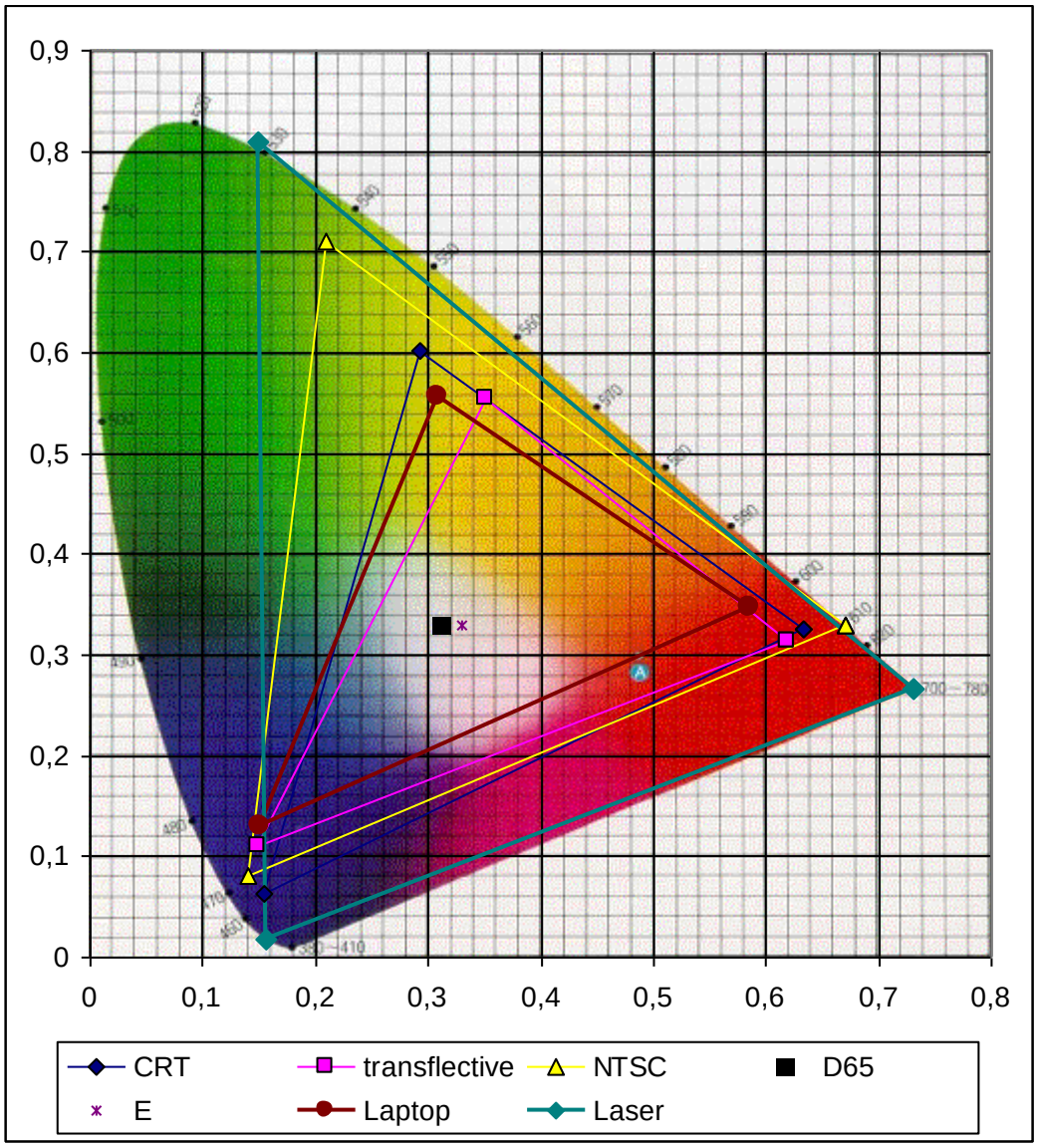

Figure 40: Comparison of transmissive/ emissive colour gamut

NTSC		CRT		Transflective		Laptop		Laser	
x	y	x	y	x	y	x	y	x	y
0.67	0.33	0.6334	0.3252	0.6188	0.3137	0.5849	0.3482	0.731	0.266
0.21	0.71	0.2934	0.6012	0.3520	0.5542	0.3083	0.5568	0.149	0.81
0.14	0.08	0.1551	0.0635	0.1497	0.1100	0.1510	0.1298	0.156	0.018
0.67	0.33	0.6334	0.3252	0.6188	0.3137	0.5849	0.3482	0.731	0.266
% NTSC	**100.0%**	**% NTSC**	**69.8%**	**% NTSC**	**52.8%**	**% NTSC**	**47.7%**	**% NTSC**	**144.5%**

Table 15: Comparison of transmissive/ emissive colour gamut

Different gamut sizes the laptop display were realised by degrading the saturation of the test images from 1 to 0 in steps of 0.1 units. The colourfulness scaling was done on a direct scale with anchor points. A test image with no colourfulness at all (saturation 0) was used as bottom anchor and a test image of the highest achievable colourfulness on the display was used as top anchor (saturation 1). The image to scale was placed in the middle and people were asked to rate it relative to the anchors on a scale of 0 to 10. The experiment was performed by 8 observers giving a total of 36 ratings on each colourfulness image. A sample screen can be seen in figure 41:

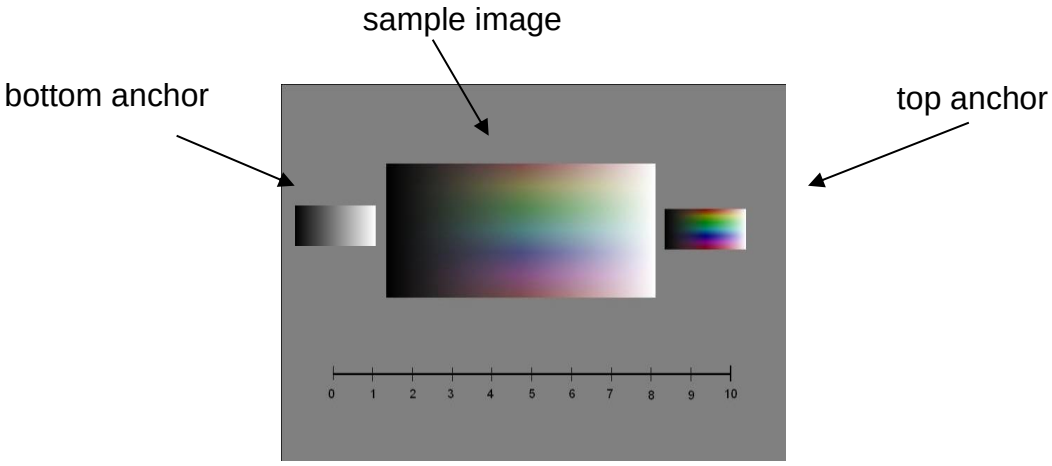

Figure 41: Sample Screen Colourfulness transmissive Display

The scaling result is given in a boxplot diagram, where the boxes indicate the borders of the upper and lower quartile of the ratings, the fat line the median and the whiskers the extent of extreme ratings:

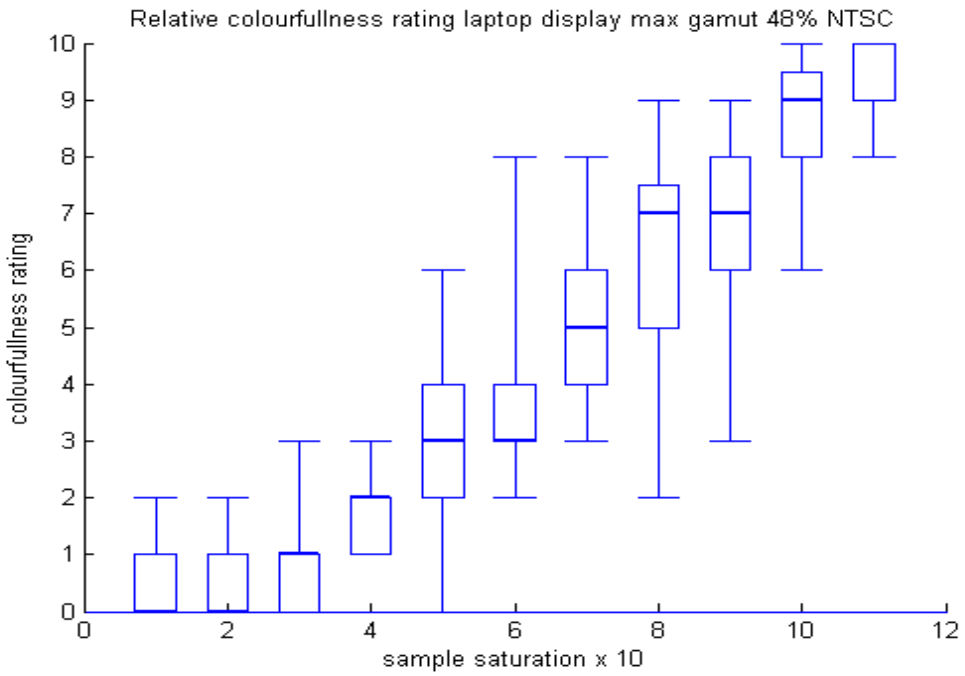

Figure 42: Colourfulness rating on laptop TFT display

The rating reveals an S-Curve relative to display saturation with noticeable rise in perceived colourfulness starting at 40% saturation and flattening starting around 90% saturation. Such an S-shape is for example typical for psychometric curves. In the next step colourfulness scaling is done on a reflective display as well.

The same kind of test image was used on the high reflective display and the reflective part of the transflectiv display (for simplicity called reflective display). Both were illuminated by a Hedler h25s Halogen spotlight with a 70cm x 70cm softbox with a colour conversion film. The colour conversion film had the effect of shifting the correlated colour temperature of the lamp from 3200 K to 5600K. Measurements of the colour gamut of both the reflective and high reflective display exhibit an even larger gamut of the reflective part of the transflective display under 5600K than calculated for D65. This stretching of the red and green primaries saturation was caused by the higher spectral radiance in these parts of the 5600K

lamp compared to D65. It is to notice that the maximum colour gamut of the reflective displays was considerably smaller than the gamut of the transmissive and the transflecitve display. The high reflective display (reflectance ~10%) had a gamut of only 9.5% NTSC for D65 while the reflective part of the transflective display had a gamut of 22% NTSC @ D65 and a considerably lower reflectance of 2.9 %.

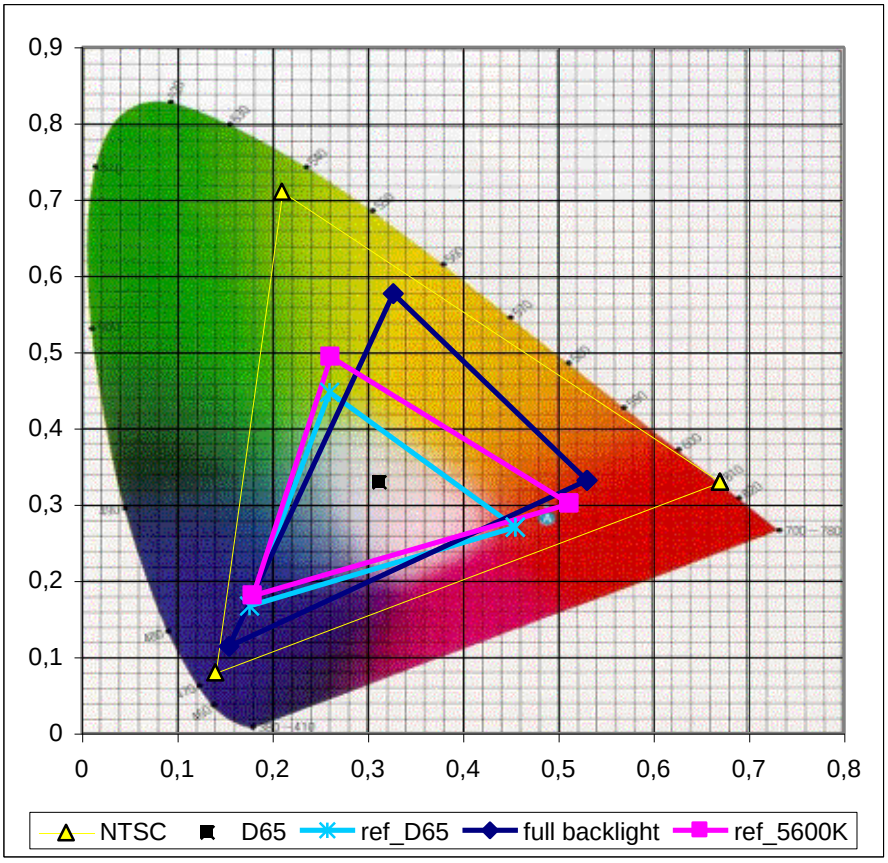

Figure 43: Colour Gamut of reflective and high reflective Display

The colourfulness scaling on the reflective was performed on the reflective display only. The reflective display was chosen because it was the one with the larger

colour gamut. The gamut sizes of the reflective display (transflective display in reflective mode) corresponding to the saturation steps are given in figure 44:

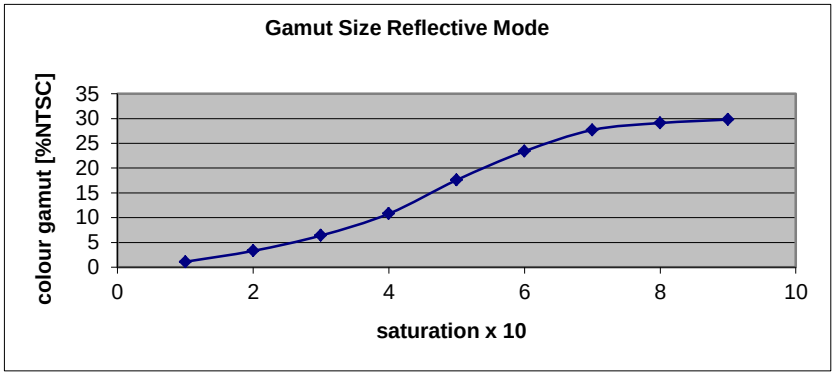

Figure 44: Gamut Sizes increasing with Image Saturation

The saturation – gamut graph of the reflective display shows an s-shaped curve comparable to the saturation – colourfulness graph of the transmissiv display in figure 42.

For the colourfulness scaling on the reflective display a paired comparison set-up was chosen. A set of colour wedge pairs covering 0.1 to 1 saturation in 0.1 steps were shown within one paired comparison sample image. The task of the observer was to indicate whether the right or the left test image was more colourful. A sample of a paired comparison image is given in figure 45:

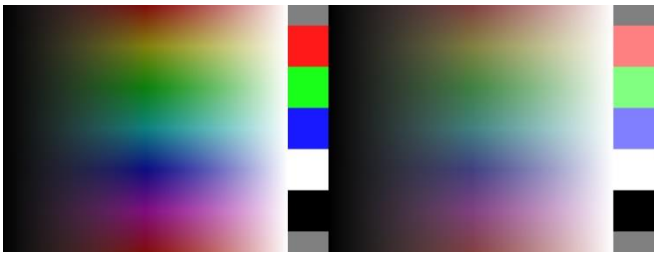

Figure 45: Paired Comparison Sample Image

A total of 6 observers took part in the paired comparison colourfulness scaling experiment. As all observers could distinguish even small steps of just 0.05 in saturation the paired comparison scaling was abandoned. Instead the next 6 observers were presented with a direct scaling task.

In the direct scaling task the right part of the image was set to the middle value of 0.5 saturation and the left part of the image varied from 0 to 1 in 0.1 steps with the exception of the value 0.5. Observers were asked to judge colourfulness of the left part of the image relative to the right part of the image on the following 7 point category degradation scale:

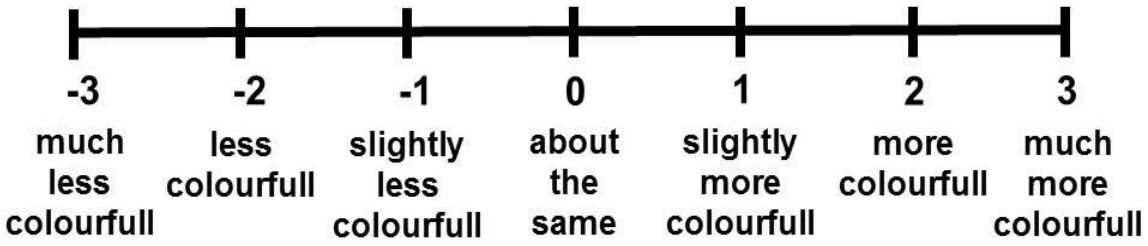

Figure 46: Degradation Category Scale on Colourfulness

When the scaling results are plotted against the colour gamut corresponding to the selected saturation values a straight line emerges, confirming the perceptual uniformity of the saturation function (figure 47):

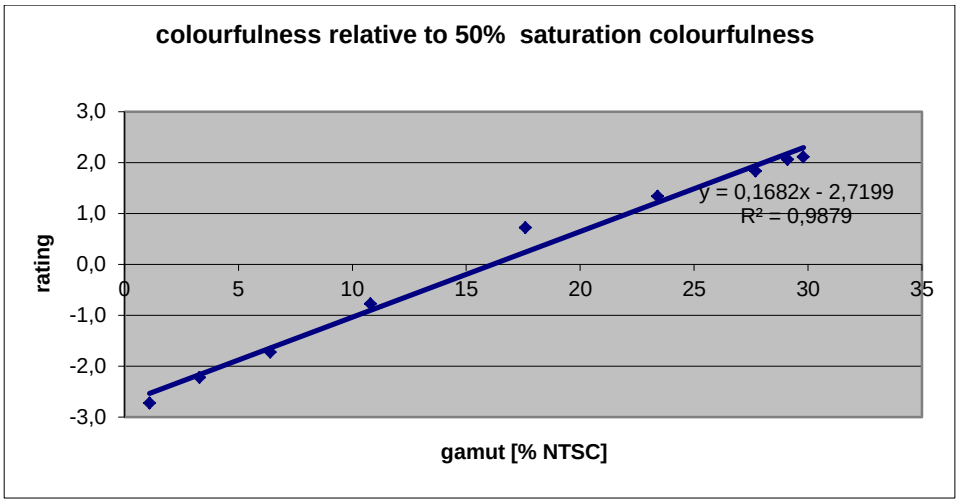

Figure 47: Result colourfulness rating vs display gamut

This means that there is a linear relationship between the physical image parameter colour gamut and the visual attribute colourfulness. Varying the colour gamut of images within one display is easiest done by varying the saturation of the image. Figure 44 showed the relationship between saturation and colour gamut. So in the following experiments colourfulness was varied in a controlled way by varying saturation of the test images. The relationship between saturation of the test images and colour gamut of the test images had been measured. The linear relationship between colour gamut and colourfulness enabled a direct comparison between colourfulness of different imaging technologies based on their colour gamut.

7.3.2 Image Quality Rating Experiments

In the image quality rating experiments the question to which extend higher colourfulness could compensate for lower brightness was investigated. This was done by comparing the image quality of an image on two displays of different

display technologies. The first pair of display technologies was the reflective part of a transflective LCD and the high reflective LCD described in chapter 5.5. The reflective part of the transflective LCD (transflective display with backlight switched off from now on called reflective display (R) for easier reading) had a relatively low reflectivity of 2.9% and a gamut of 22% NTSC. The high reflective display had a reflectivity of 10%, but a colour gamut of only 9.5% NTSC. Because of the high brightness and low gamut the high reflective display was chosen as reference. The first series of comparisons were varying saturations on the reflective display (R) versus full saturation on the high reflective display (HR). In the second round the reference image on the high reflective display was reduced to a saturation of 0.7 (0.7 HR).

The map image in figure 48 was chosen because it really had red, green and blue parts in it:

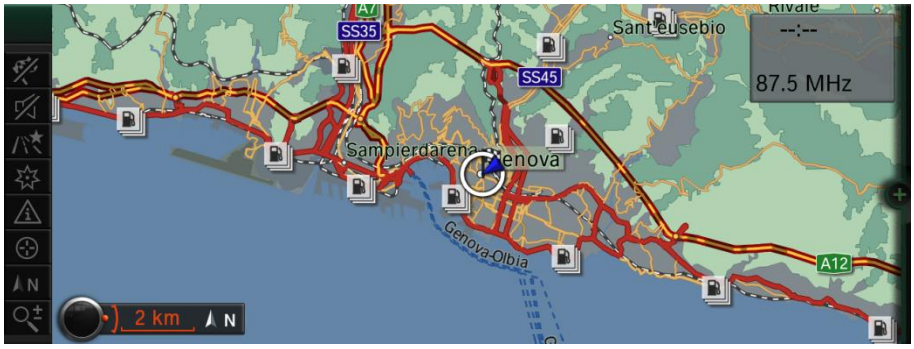

Figure 48: Map Image for colourfulness and image quality rating

A total of 10 map images with maximum saturation ranging from 0.1 to 1 were prepared. First these were shown in random order on the reflective display and observers were asked to rate the perceived image quality compared to reference

images on the high reflective display. The first reference image was a full gamut and the second reference image was a 0.7 saturation gamut image on the high reflective display. The rating was done on a 5 point scale going from 1 considerably worse, 2 worse, 3 about the same, 4 better to 5 considerably better:

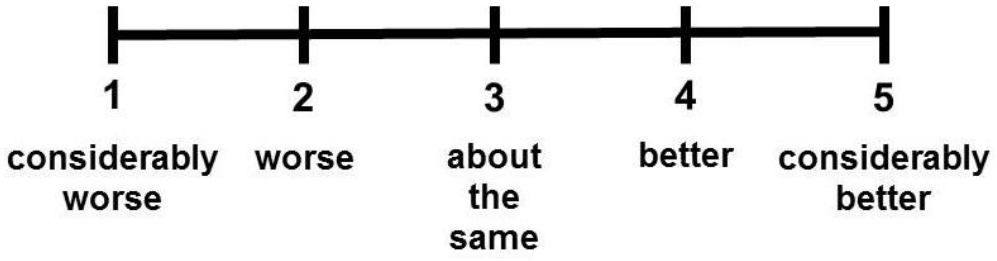

Figure 49: Rating scale for image quality rating experiments

The scaling results are given in figure 50:

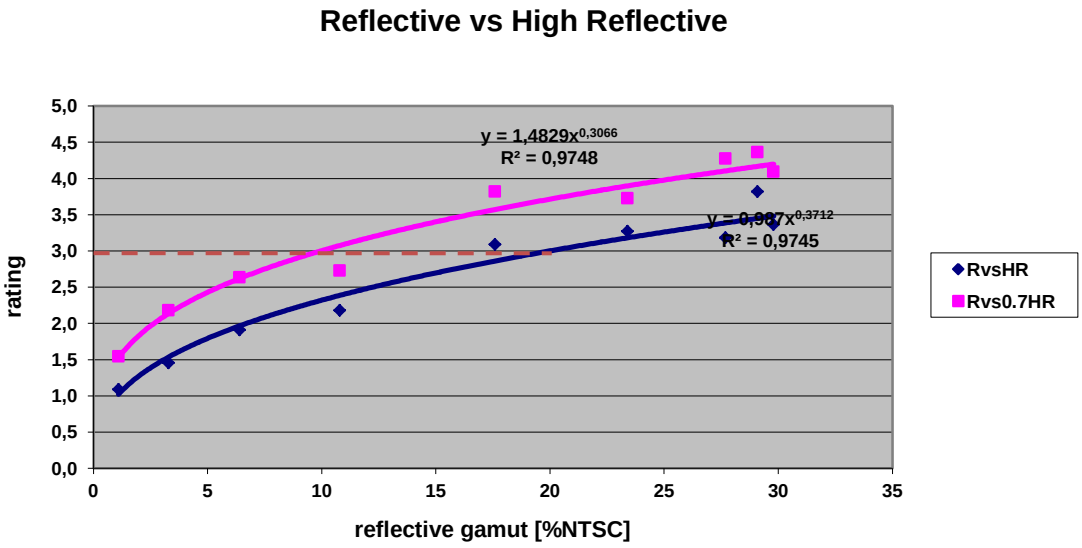

Figure 50: Image Quality Rating Reflective versus High Reflective

Note that for the full saturation HR reference image (gamut 9.5%) an image with a colour gamut of around 20% on the darker reflective display (R) was needed to achieve the same perceived image quality as on the HR reference display (rating 3). This was nearly twice the gamut of the high reflective gamut. The high score of the high reflective display in spite of its rather small colour gamut was due to the positive impact of the higher brightness of the high reflective display. The reflectivity of the HR display was 10% compared to 2.9% for the reflective display (R), so roughly speaking in this case colour gamut had to be doubled in order to compensate for a 3 times higher brightness.

The next comparisons were done with the backlight on i.e. in transflective mode.. The full gamut of the transflective display was 53% NTSC and luminance of the map image was 273 cd/m². The reference images were the full gamut image on the high reflective display (HR: 9.5% NTSC, L_{max} 200cd/m²) and a full gamut image of the laptop (transmissive display (TR: 48% NTSC, L_{max} 188 cd/m^2)). The scaling results are given figure 51:

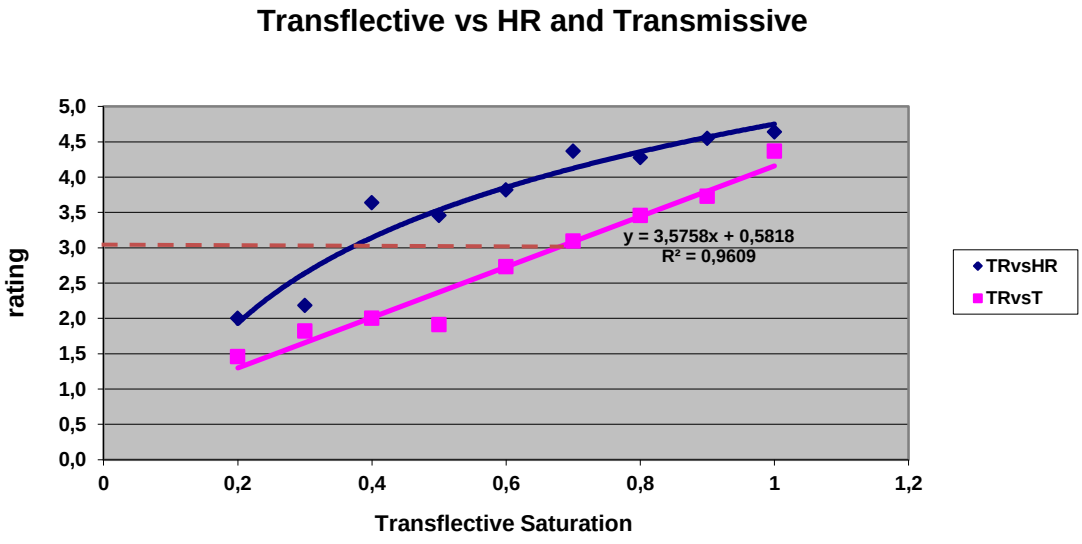

Figure 51: Scaling Result Transflective Display with full backlight

Compared to the high reflective display (HR) the display in transflective mode (TR) was judged better already from a low colour gamut of 6.4% NTSC with the rating curve going in saturation from a colour gamut of 20%. As the transflective display was brighter than the high reflective display, the high reflectivity was no longer able to compensate for the small colour gamut of the high reflective display.

When comparing the transflective display (TR) to the laptop display (T) outperformance started at a gamut of 23.4 % NTSC corresponding to a saturation of 0.7 As the gamut of the laptop was about 48% NTSC, which was very close to the full gamut of the transflective display of 53% NTSC, there are other parameters outperforming colour gamut. In the case of the transflective display this was the highly appreciated sharpness of the display and its brightness even though the colours on the laptop were judged to be more natural especially in the greens and blues.

7.4 Discussion Colourfulness Experiments

The image quality experiments showed that independent of display technology a higher colourfulness can compensate for lower brightness. In the experiment on the reflective technologies image quality the high reflective display which had about 3 times the reflectivity of the reflective display was judged to be of the same image quality, when the gamut of the reflective display was about twice the gamut of the high reflective display.

The transflective display was the brighter than the high reflective and transmissive display it was compared to. In order to be considered to have the same image quality as the other displays, the transflective display needed only a fraction of its

colour gamut. Interesting is that even though the brightness of the reflective display and the transmissiv display are comparable (L_{max} 200 cd/m² vs. L_{max} 188 cd/m²), the transflective display is judged better than the high reflective display (HR) already at a colour gamut of 6.4 % NTSC (ca. 70% of HR gamut) and better than the transmissive display (T) at a colour gamut of 23.4% NTSC (ca. 50% of T gamut). Even though this experiment showed as well, that higher colourfulness can compensate for lower brightness, it had been expected, that even higher colour gamut would have been necessary on the transflective display to achieve the same perceived image quality as on the transmissive and reflective display. The main reason observers reported for the stronger than expected perceived image quality ratings of the transflective display was very high sharpness of the transflective display compared to the other displays. This result strongly suggest, that even though perceived contrast, blackness, brightness and colourfulness are all relevant attributes for image quality assessment, an image quality rating across display technologies is not complete unless sharpness of the image is taken into account as well.

8 Conclusions and Future Work

The aim of this thesis was to link perceived image quality to physical display parameters. This was done for the physical display parameters black and white luminance, contrast and colour gamut. As the major part of this research was performed at BMW Group, a further aspect of this thesis was to describe environmental conditions affecting perceived image quality of automotive displays. In this final chapter the subchapter conclusions summarises what has been actually achieved within this thesis and the subchapter future work suggests the next steps future research could take on the way to formulate an image quality model for automotive displays.

8.1 Conclusions

In this thesis perceived image quality of automotive displays was investigated based on the framework of the image quality circle by Engeldrum (2000). Even though there are models available, which predict readability of a display under lighting conditions relevant to automotive applications (Adrian 1987, Dreyer 2007, Silverstein 1996, BAE Systems 2001) these models are limited in predicting perceived image quality. A display which scores higher in terms of readability can be expected to score higher in terms of perceived image quality as well. Especially in the automotive environment, where eyes-off-road-time has to be minimised as much as possible for safety reasons, guaranteeing readability under all environmental conditions is a must. However, once readability is given, the next level to rate the performance of a display is perceived image quality. According to the image quality circle observers form their image quality rating by weighting the

visual attributes they perceive. Visual attributes describing displays are for example colourfulness, brightness, sharpness or perceived contrast (“contrastness”). The functions translating physical image parameters which can be measured into perceived visual attributes are called visual algorithms.

In this thesis the visual attributes brightness, blackness, contrastness as functions of the DICOM JND scale, as well as colourfulness as a function of display gamut, have been investigated in their impact on image quality rating of automotive displays. It has been shown in chapter 7.3.1 that the visual attribute colourfulness rises linearly with the physical attribute colour gamut. In chapter 7.3.2 it has been shown that higher colourfulness can compensate for lower brightness in image quality rating. In the experiments on brightness – contrastness preference one result was that higher contrasts were preferred as long as the display does not get uncomfortably bright. Glare had a detrimental effect on perceived image quality and the luminance level causing glare strongly depended on adaptation of the observer. The other detrimental effect showing in the brightness-contrastness experiments was found at the bottom end of the luminance range of a display. The perceived “blackness” of black was shown to influence image quality rating of observers. A symbol of the same perceived contrast (same ΔJND) was rated better when the background was perceived as a deep black than a symbol where the background luminance was slightly higher and the background was perceived as degraded black. In the preference for brightness-contrastness combinations in chapter 7.1.4 in dark environment and in the office environment (chapter 7.1.5) perceived contrasts covering only a small percentage of the dynamic range were ranked best when using the middle of the JND range of the display (presentations around mid-grey). This means for simple information displays which only have to

convey binary images relatively cheap monochrome displays with a low dynamic range could be used successfully as long as background luminance is not perceived as black and foreground luminance does not cause discomfort glare. The office environment experiment showed in principle the same behaviour with the difference that the region perceived as mid grey was shifted to higher luminance levels.

Already the short pilot on the brightness- “contrastness” experiment under office lighting conditions showed the importance of adaptation on perceived appearance. Even though a 200 lux illuminance perpendicular to the screen did not evoke high adaptation luminance, the rise in adaptation luminance was sufficient to shift the display’s luminance out of the discomfort glare region. When considering the much higher adaptation luminance and display illuminance values encountered in automotive viewing scenarios a severe influence on display appearance and quality rating is to be expected.

The strong reactions to discomfort glare and the importance of the blackness of black as long as it is recognized as black are both parameters not covered by typical readability models. This supports the view that there is more to image quality then sheer task performance. For the application of interior lighting Boyce (2003, p.191) goes even further: *“Eliminating visual discomfort is not a recipe for good quality lighting. Rather, it is a recipe for eliminating bad quality lighting and replacing it with indifferent lighting. This would be no small achievement.”*

8.2 Future Work

In this thesis the image quality scaling experiments were done with a limited set of visual attributes. Most experiments were performed in the easy to control and reproduce environment of a dark room or with relatively moderate ambient illumination. The next logical steps would be to increase the number of visual attributes included in the image quality scaling experiments and to perform image quality scaling experiments under ambient lighting conditions which are representative for driving at daylight.

8.2.1 Suggestions on Visual Attributes

The visual attribute investigated in this thesis were colourfulness, brightness, blackness and contrastness. This makes four visual attributes which contribute to image (display) quality. Further visual attributes mentioned in chapter 3.2.2 contributing to image/ display quality and not investigated in the experimental part are sharpness, homogeneity and glossiness.

The visual attribute homogeneity is in fact part of current specifications of automotive displays. A rather simple formula for acceptable deviation from the average display luminance does ensure a high standard of homogeneity within automotive displays. For this reason it is considered to be feasible to omit lengthy homogeneity scaling as long as displays meeting the internal homogeneity standard are used in image quality rating experiments.

Glossiness is function of the reflective properties of a display as expressed in BRDF or BTDF functions described in chapter 2.2.2. These measurements are relatively time-consuming and the preparation of display samples with exactly tuned BRDF functions would prove even more complicated. In the experimental

part no observer mentioned glossiness as a visual attribute influencing his image quality rating. For this reason glossiness is not judged as the most important attribute to be investigated in future experiments.

However, in experiment 7.3.3 subjects explicitly mentioned sharpness as one reason for their image quality rating even though the intention was to investigate the influence of brightness and colourfulness on image quality rating. This leads to the conclusion that sharpness is a visual attribute with strong impact on perceived image quality and should be considered when rating image quality of display technologies.

Especially in the research for TV applications a number of researchers have already investigated sharpness (Liu et. al. 2004, Xia et. al. 2003, Heynderickx et. al. 2002). In automotive applications sharpness becomes for example an issue for LCD displays at very low display temperatures due to stronger differences in the switching times of individual grey levels. So far sharpness is not a point generally found in specifications for automotive displays, but it is felt that including sharpness in image quality rating would give more accurate benchmarking results when comparing display technologies.

8.2.2 Impact of the Automotive Environment

Fragments of Image Quality Models based on assessment of visual attributes do already exist (Heynderickxs 2005, Quin 2006; Bech 1996, Pointer 2003). Some of them do take even more aspects into account than possible in this PhD project. However, all of these approaches to image quality are based on applications in less demanding surroundings, be it television or typical office applications. Future

research could be to include the evaluation of the influence the demanding automotive environment has on the perception of the physical stimuli and the quality rating. Figure 52 shows where the environment influences the Image – Quality –Circle:

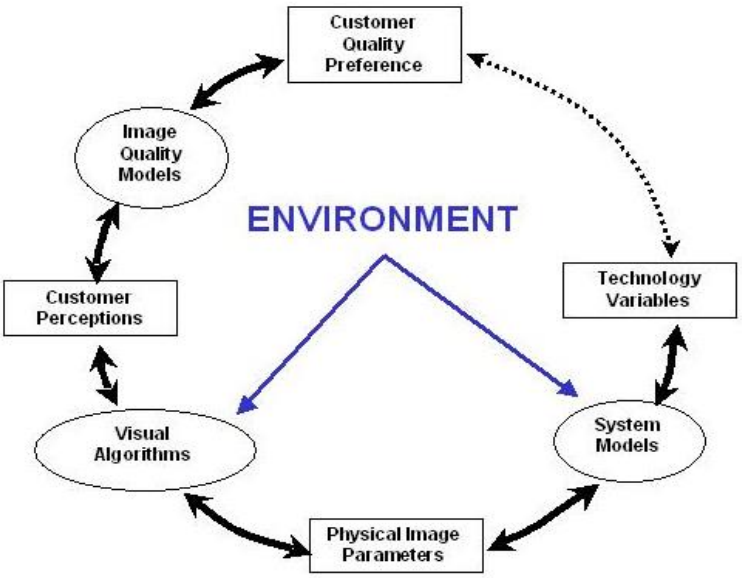

Figure 52: Environmental influences to the Image-Quality-Circle

First of all the environment has an impact on the stimuli which can be produced by a given display technology. For display applications the most important environmental factor is ambient light. Other factors like temperature, vibration etc. have impacts as well, but should not be included in a display model instead of an image quality model.

The influence of ambient light on how well physical stimuli can be perceived is an important aspect in all readability models mentioned in section 2.2. The most prominent factor here is the observer's adaptation luminance L_a; which is part of all

models. How L_a is defined, already differs between the three models mentioned. The PJND model uses the background luminance of the displayed information normalised to 10,000 cd/m². The TTV model uses a variety of field of forward view (FFOV) values defined for different applications. For the Visiblity Level (VL) model Dreyer (2007) integrates over the scene luminance within a 30°field of view. One of the first tasks in investigating perceived image quality under ambient lighting conditions relevant to driving should be to decide on the most sensible definition of key driver adaptation levels and corresponding illumination on the display.

None of these models deals with chromatic adaptation, which can become important in colour perception. To which degree chromatic adaptation becomes an issue in a driving scenario could be investigated by performing image quality scaling experiments under illumination with 2700K and with 6500K, if no significant differences in image quality rating are achieved between these 2 scenarios, chromatic adaptation could be ignored.

In case the colour temperature of the ambient illumination and the adapting white point prove to be important, the main experiment on the impact of the lighting scenario should take place with light sources carefully adjusted to relevant daylight colour temperatures. If not, the only focus should be on realising relevant adaptation luminance and display illuminance conditions. Image quality ratings might be performed under a reference condition of diffused illuminance of 3000 lux and an adaptation luminance of 3000 cd/m² representing Silverstein's low ambient daylight scenario (Silverstein 1996). In the test set-up adaptation luminance and display illuminance should cover a substantial range of possible automotive conditions and should be controlled independently by the investigator. Observers could be subjected to a range of combinations of adaptation luminance and

display illuminance. Within each of these scenarios their task would be to give image quality ratings for all investigated display technologies. The result would be perceptual performance curves for all investigated display technology types over the tested range of automotive lighting conditions.

Once all named influences of the automotive environment on attribute perception and image quality rating would be established and fitted into a model of automotive image quality the model should be tested under real world conditions. This might be done by asking observers to assess image quality of a display under the reference conditions of 3000 lux illuminance and 3000 cd/m² adaptation luminance and in a real car outdoors. For the model to be valid the image quality rating under the measured adaptation luminance and display illuminance should fit the models predictions for the given parameters.

APPENDIX

A1 The Barten Model – A Contrast Sensitivity Function

Barten (1992) models the contrast sensitivity of the human eye. Input parameters are luminance and display size on the basis of internal noise of the visual system. The Barten Model gives a global description of the optical MTF of the human eye. Generally contrast sensitivity is defined as inverse of the threshold contrast i.e. a contrast which is just perceptible. Barten assumes that this threshold contrast is completely determined by noise, be it internal noise of the human visual system or possible external noise which may be present in the viewed image.

The influence of external noise on the threshold contrast M_t of a sinusoidal grating pattern is given by:

$$M_t^{'} = \sqrt{M_t^{\,2} + (kM_n)^2})$$

Equation 29: Threshold contrast in presence of external noise

M_t' is the reduced threshold contrast in the presence of external noise, M_t is the threshold contrast without noise, M_n is the modulation depth of the noise wave components and k is a dimensionless constant. The constant c appears as well in formulas defining the signal to noise ratio of just observable objects. For the sinusoidal patterns used in the Barten model values ranging from 2.6 to 5.2 were found. Cognitive aspects like advanced knowledge of the observed pattern, the

individual observer and methodological aspects like the method used to determine the threshold influence the actual value of the constant k.

The modulation depth M_n of the noise wave components is calculated from the spatial and temporal spectral density Φ_n of the noise. This is where the spatial and temporal integration capacity of the human visual system comes into the game. For static images the integration time of the eye T is in the order of 0.1 sec. In the spatial dimensions X and Y the limits are the maximum angular size and the maximum number of cycles over which the eye can integrate information. The effect of these limits on the spatial dimension X (and Y respectively) can be described as follows:

$$X = \frac{1}{\sqrt{\frac{1}{X_0{}^2} + \frac{1}{X_e{}^2} + (\frac{u}{N_e})^2}}$$

Equation 30: Spatial dimensions of a considered picture

X_0 is the angular size of the observed picture, X_e is the maximum angular size and N_e is the maximum number of cycles over which the eye can integrate the information while u is the angular spatial frequency.

In short the integration angle X is equal to the smallest of the three components object size, maximum integration angle and visual angle determined by the number of cycles.

Barten chose an average N_e of 15 cycles from the maximum number of cycles over which the eye can integrate from the numbers reported in literature. For the maximum integration angle X_e Barten uses a value of 12°.

While the dashed lines represent experimental data, the continuous lines are derived from the above equations with T = 0.1 sec, X_e =12°, N_e = 15 and k = 3.1 while for Mt the data measured without noise was used.

Barten applies the equations derived for external noise for internal noise as well. The most important sources for internal noise of the human visual system after Barten are fluctuations in the photon current arriving at the photoreceptors and neural noise generated in the nerve system. Furthermore Barten takes into account that the information is filtered by the optical MTF of the eye before reaching the retina.

Photon noise is caused by statistical fluctuations in the photon current which depends on the luminance of the object as well as the pupil diameter. The photon flux J is given by:

$$J = pI$$

Equation 31: Photon flux J passing through the cornea

The illuminance I of the eye expressed in trolands is dependent on the luminance of the object as well as the pupil diameter which itself is again dependent on

luminance. The variation of the natural pupil can be fitted with a hyperbolic tangent. That means the illuminance of the eye I is defined by the following two equations:

$$I = \frac{\pi}{4} d^2 L$$

Equation 32: Illuminance of the eye in trolands

$$d = 4.6 - 2.8 \tanh(0.4 \log(L / L_0)) \quad [mm]$$

Equation 33: Pupil diameter

For chromatic light under photopic viewing conditions the value of the numerical constant p in equation 3 is given by:

$$p = 0.6270 \frac{\int P(\lambda) V(\lambda) \lambda d\lambda}{\int P(\lambda) V(\lambda) d\lambda} \quad [photons/(td \cdot \sec \cdot arc\min^2)]$$

Equation 34: Constant p linking illuminance of the eye I to photon flux J

P(λ) is the spectral energy distribution of the light source and V(λ) the relative spectral sensitivity of the human visual system for a wavelength λ in nm. For scotopic conditions a factor of 0.2442 and V'(λ) have to be used instead.

Not all photons entering the eye excite the photoreceptors on the retina. In fact one has to calculate with a quantum efficiency η of a few percent or less. So the spectral density of the noise caused by variations of the photo current is:

$$\Phi_{ph} = \frac{1}{\eta J}$$

Equation 35: Spectral density of photon noise

Neural noise is noise generated in the neural system. It is well known that the contrast sensitivity of the human visual system is attenuated at low spatial frequencies and exhibits a nearly linear increase with spatial frequency at the mounting slope of the spatial contrast sensitivity curve. For this reason neural noise can't be assumed to be simple white noise i.e. independent of spatial frequency. The existing dependence on spatial frequency of the contrast sensitivity can be modelled by the following equation:

$$\Phi_{neu} = \Phi_0 \Big/ (1 - \exp(\frac{-u^2}{u_0{}^2}))$$

Equation 36: Characteristic behaviour of neural noise

Φ_0 is the noise density at high spatial frequencies and u_0 the spatial frequency below which the attenuation of the contrast sensitivity occurs. Generally this frequency is reported to be around 8 cycles per degree.

Barten interprets the above equation as follows: Φ_0 is the real neural noise and the attenuation is caused by a filtering process in the early stage of the neural processing. A very likely candidate for this kind of filtering is the lateral inhibition caused by the ganglion cells which is known to suppress low spatial frequencies. A common mathematical description of the lateral inhibition process is the subtraction of a spatially low pass filtered signal from the original signal. When the MTF of the low pass filter is given by F(u) the MTF of the inhibition process can be expressed as 1-F(u). This leads to a more general form of the equation for neural noise:

$$\Phi_{neu} = \Phi_0 \big/ (1 - F(u))^2$$

Equation 37: Neural noise at the entrance of the visual system

The combined equations for photon and neural noise give the total internal noise:

$$\Phi_{\text{int}} = 1 \big/ (\eta p I) + \Phi_0 \big/ (1 - F(u))^2$$

Equation 38: Total internal noise of the human visual system

Barten uses a value of 3E-8 sec deg² for the original neuron noise Φ_0 which sets a limit to the contrast sensitivity at high luminance levels.

Apart from the limited amount of integration cycles the optical MTF of the eye is another reason for decreasing contrast sensitivity at higher spatial frequencies. Effects like spherical and chromatic aberration of the lens, straylight from optical

media, diffusion in the retina as wells as the discrete structure of the rods and cones contribute to the optical MTF of the eye. Barten assumes that the total effect of all these individual contributions can be expressed by a simple Gaussian MTF given by:

$$M_{opt}(u) = \exp(-\pi^2 \sigma^2 u^2)$$

Equation 39: Optical MTF of the human eye

The radial standard deviation of optical point spread function resulting from the convolution of the named contributions is given by σ. Because of the spherical aberration of the eye lens at large pupil sizes the value of σ increases with the third power of the pupil diameter:

$$\sigma = \sqrt{{\sigma_0}^2 + (C_{sph} d^3)^2}$$

Equation 40: Influence of pupil diameter on radial standard deviation

The value of σ at small pupil sizes is given by σ_0 while C_{sph} is a constant describing the spherical aberration effect. Barten expects the value of C_{sph} to be reasonably observer independent to justify the use of an experimentally derived value of 0.006 arcmin/mm³ as a fixed constant in the model. The same cited experiments which suggest a C_{sph} of 0.006 arcmin/mm³ indicate a σ_0 of about 1 arcmin which is however suspected to vary from observer to observer.

The total Barten Model merges all the previously explained equations into the following overall formula for the contrast sensitivity function:

$$\frac{1}{M_t} = \frac{1}{k} \frac{\sqrt{T/2}}{\sqrt{\left(1/(\eta p I) + \Phi_0/(1-F(u))^2\right)\left(1/X_0^2 + 1/X_e^2 + (u/N_e)^2\right)}} M_{opt}(u)$$

Equation 41: Complete contrast sensitivity function of the Barten Model

Barten proved a good match to experimental data published by other researchers by adapting the remaining constants between the following values:

constant	range	dimension
p	339-400	Photons/td sec
k	2.6-4.7	
ηη	0.5-10	%
σσ$_0$	0.60-1.70	arcmin

Table 16: Ranges for adaptable contestants of the Barten Model

REFERENCES

Adrian, W. 1989. Visibility of targets: Model for calculation. *Lighting Research and Technology December.* 21(4): pp.181-188.

Aleksander, D.K. 2000. *Daylight Modelling: the sky dome.* [photograph in online-pdf]. Welsh School of Architecture, Cardiff University. Available from: http://www.cardiff.ac.uk/archi/skydome .[Accessed 15 June 2006].

BAE Systems 2001. *Generic Lighting Standard.* Report No.BAE-FSE-R-RES-CR-05028. Warton: BAE Systems.

Barten, P. G. J. (1992). Physical Model for the Contrast Sensitivity of the Human Eye, in: *Human Vision, Visual Processing and Digital Display III*. SPIE. 1666: pp.57-72.

Bech, S. et al. 1996. The rapid perceptual image description method. *SPIE Proceedings.* 2657: pp.317–328.

BMW Group. 2007a. *Motor Vehicles Night Design Interior.* GS 95010. Munich: unpublished.

BMW Group. 2007b. *Electrical/Electronic Assemblies in Motor Vehicles –Part 4 Klima.* GS 95003. Munich: unpublished.

BMW Group. 2008. *View from the driver's perspective.* [Rendering]. Unpublished.

Boyce, P.R. 2003. *Human Factors in Lighting.* 2nd ed. London: Taylor and Francis.

Brennesholtz, M. S. 2006. Expanded-Color-Gamut Displays: Part 1: Colorimetry for Video Signals and Display Systems, *Information Display,* 9: pp.24-29.

Cartwright, C. 2007. Readability of Displays at High Ambient Lighting Conditions, Paper presented at: *SID Display Week.* October 2007.

Chen, L.C. and Kuo, C.C. 2008. Automatic TFT-LCD mura defect inspection using discrete cosine transform-based background filtering and 'just noticeable difference' quantification strategies. *Measurement Science and Technology.* 19(1): 015507. Available from: doi:10.1088/0957-0233/19/1/015507. [Accessed 2 February 2008].

Commission International de l'Eclairage. 1992. CIE 095-1992, *Contrast and Visibility.* Vienna: CIE.

Commission International de l'Eclairage. 2001. CIE 142-2001, *Improvement to industrial colour difference evaluation.* Vienna: CIE.

Commission International de l'Eclairage. 2004. *Colorimetry,* CIE 015:2004, CIE, Vienna.

Commission International de l'Eclairage. 2004. ISO 15469:2004 (E)/CIE S 011/E, 2003*Spatial Distribution of Daylight – CIE Standard General Sky.* Vienna: CIE.

Commission International de l'Eclairage. 2008. *Activity Report Division 1 Vision and Color* [online]. Available from: http://www.cie.co.at/div1/ActReps/D1ActivityReport08.pdf. [Accessed 5 March 2009]

Commission International de l'Eclairage. 2011. CIE S 017/E:2011, *ILV : International Lighting Vocabulary.* Vienna: CIE.

Daiwa House (2006). *Mirror cabinet artificial sky laboratory.* [online image], Nara, Japan, Daiwa House Industry Co., LTD, Central Research Laboratory, Available from: www.daiwahous.co.jp/lab/en/equipments/equip06.html [Accessed 15 August 2006].

Deutsches Institut für Normung. 1982. DIN 67 530, *Reflektometer als Hilfsmittel zur Glanzbeurteilung an ebenen Anstrich- und Kunststoff-Oberflächen*. Berlin: Beuth Verlag.

Deutsches Institut für Normung. 2001. DIN 6176, *Farbmetrische Bestimmung von Farbabständen bei Körperfarben nach der DIN99-Formel*. Berlin: Beuth Verlag.

U.S. Department of Transportation. 1995. Federal Highway Administration, FHWA-RD-94-089: *Suggested Procedures and Acceptance Limits for Assessing the Safety and Ease of Use of Driver Information Systems*. Georgetown Pike McLean (VA): DOT.

Dreyer, D. 2007. *Kontrastschwellensimulation für Sichtbarkeitsuntersuchungen an Displays.* [PhD thesis]. Technische Universität München, Fakultät für Maschinenwesen.

Eckert, M. 1993. *Lichttechnik und optische Wahrnehmungssicherheit im Straßenverkehr.* München: Verlag Technik Berlin.

E Ink Corporation. 2002. *Active matrix electrophoretic information display for high performance mobile devices.* Cambridge (MA): E Ink Corporation.

Eloholma, M. et. al. 2005. Mesopic models – from brightness matching to visual performance in night-time driving: a review. *Lighting Research and Technology.* 37(2): pp.155-175.

Littlefair, P. 1985. The luminous efficacy of daylight: a review. *Lighting Research and Technology. 17(4): pp.162-182.*

Engeldrum, P.G. 2000. *Psychometric Scaling: A toolkit for Imaging Systems Development.* Winchester: Imcotek Press,.

Fairchild, M.D. 2005. *Color Appearance Models, 2nd Edition.* Chichster: John Wiley & Sons Ltd.

Goldstein, E.B. 2008. *Wahrnehmungspsychologie Der Grundkurs.* 7th ed. Irtl, H. ed. German edition. Heidelberg: Springer.

gretagmacbeth 2005. *Certificate of Performance SpectraLight Viewing System*., New Windsor (NY): gretagmacbeth.

Heimrath, M. 2000. Invited Paper: Roadmap to Fulfil the FPD Need of Future Cars. *SID 00 Digest.* pp.1149-1151.

Heynderickx, I., Matenbroek, E. and Bech, S. 2002. Sharpness as perceived on a CRT. In: *Proceedings of the International Display Workshop IDW' 02, 4-6 December 2002,* Hiroshima, Japan*.* Malden (MA): Wiley. p.1267.

Heynderickx, I. and Langendijk, E.H.A. 2005. Image Quality Comparison of PDP, LCD, CRT and LCoS Projection. *SID 05 Digest*. pp.1502-1505.

Hopkinson, R.G. and Collins, J.B. 1970. *The Ergonomics of Lighting*. London: McDonald & Co.

International Organization for Standardization. 1994. ISO 2813:1994(E), *Paints and varnishes – Determination of specular gloss of non-metallic paint films at 20°, 60°, and 85*°. Geneva: ISO.

International Telecommunication Union. 2002. REC. 709-5; ITU-R, *Parameter values for the HDTV standard for production and international programme exchange*. Geneva: ITU-R.

International Organization for Standardization. 2003. ISO 15008:2003, *Road vehicles- Ergonomic aspects of transport information and control systems. Specifications and compliance procedures for in-vehicle visual presentation*. Berlin: ISO.

International Organization for Standardization. 2005. ISO 23539:2005(E)/CIE S 010/E:2004, *Photometry - The CIE System of Physical Photometry.* Berlin: ISO.

International Organization for Standardization. 2007. ISO 11664-2:2007(E)/CIE S 014-2/E:2006, *Colorimetry — Part 2: CIE Standard Illuminants for Colorimetry.* Berlin: ISO.

Jones, J.C. 2007. Invited Paper: The Zenithal Bistable Dievice: From Concept to consumer. *SID 07 Digest.* pp.1347-1350.

Judd, D.B. and Wyszecki, G. 1975. *Color in business, science and industry.* 3rd ed. New York: Wiley.

Kelly, E.F. 2002. Sensitivity of Display Reflection Measurements to Apparatus Geometry. *SID 02 Digest*. pp.140-143.

Liu, L.; Xia, J. and Heynderickx, I. 2004. Study of Sharpness Threshold for Different Cultural Backgrounds. *ASID 04 DIGEST.* p.274.

MacAdam, D.L. 1942. Visual sensitivities to color differences in daylight. *Journal of the Optical Society of America.* 32: pp.247-274.

MacAdam, D.L. 1942. Specification of small color differences in daylight, *Journal of the Optical Society of America.* 33: pp.18-26.

Muck, E. and Bodmann, H.W. 1961. Die Bedeutung des Beleuchtungsniveaus bei praktischer Sehtätigkeit. *Lichttechik*. 13: pp.502-507.

National Electrical Manufacturers Association. 2003. PS 3.14-2003, *Digital Imaging and Communications in Medicine (DICOM) – Part 14: Grayscale Standard Display Function*. Rosslyn (VA): NEMA.

National Physical Laboratory. 2006. *Good Practice Guide for the Measurement of Gloss.* Teddington: NPL.

Poynton, C. 2007. Wide-Gamut Displays. *Information Display.* 7: pp.10-15.

Qualcomm MEMs Technologies Inc. 2009. *IMOD Technology Overview*. [White paper]. Available from: www.qualcomm.com/sites/default/files/uploads/imod-tech-overview-06-2009.pdf [Accessed September 2009].

Quin, S.L. et al. 2006. Perceptually Relevant Characterization of LCD Viewing Angle. *SID 06 DIGEST*. pp.1320-1323.

Robins, C.L. 1986. *Daylighting Design & Analysis.* New York: Van Nostrand Reinhold Company.

Schweigert, M. 2003. *Fahrerblickverhalten und Nebenaufgaben.* [PhD thesis]. Technische Universität München, Fakultät für Maschinenwesen.

Sharpe et al. 2006. A usability metric for displays in challenging environments, Paper presented at: *International Display Workshop 06, Ostsu, Japan, December 2006*.

Silverstein, L.D. 1996. *Display visibility in dynamic lighting environments: Impact on the design of portable and vehicular displays*, Presentation to Motorola University, 1997.

Sony Electronics Inc. 2006. *PRS-500 Portable Reader System.* [Marketing Specifications]. Available from: https://docs.sony.com/release/specs/PRS500_mksp.pdf [Accessed 08. January 2013].

Stevens, S.S. 1959. On the measurement of sensation. *Acta Psychologica.* 15: pp.91-101.

Vassie, K. 1998. Specification and assessment of visual aspects of cockpit displays, *SID 1998 Digest.* 29(1): pp.1199-1203.

Video Electronics Standards Association. 2001. *Flat Panel Display Measurements Standard Version 2.0.* Milpitas (CA): VESA Display Metrology Committee. pp.271-278.

Walkey, H.C. et. al. 2006. Characterising mesopic spectral sensitivity from reaction times. *Vision Research*, 46: pp.4232-4243.

Xia, J., Yin, H.C. and Heynderickx, I. 2006. Preferred balance between luminance and color gamut in mobile displays. *Journal of the SID*. 14(10): pp.873-881.

Xia, J., Liu Y. and Yin, H.C. 2003. Relation between Measured Spot Profile and Perceived Sharpness. IDW 03: *Proceedings of the International Display Workshop, 3-5 December 2003,* Fukuoka, Japan. Malden (MA): Wiley. p.1475.

Xu, M. and Yang, D. 1999. Optical Properties of the Gray-Sxale States of Cholesteric Reflective Displays. *SID 99 Digest.* pp.950-953.