

Physical and sensorial image descriptions

Citation for published version (APA):

Hamberg, R. (1994). *Physical and sensorial image descriptions*. (IPO-Rapport; Vol. 1013). Instituut voor Perceptie Onderzoek (IPO).

Document status and date:

Published: 21/10/1994

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Institute for Perception Research
P.O. Box 513 - 5600 MB Eindhoven

RH/rh 94/09
21.10.1994

Rapport no. 1013

Physical and Sensorial
Image Descriptions

R. Hamberg

Preface

This report is completely embedded in the ADONIS project, which is a collaboration between Bang & Olufsen, who initiated it in the form as it is, Philips Consumer Electronics, who is the coordinator of this project, Philips Components, and IPO; ADONIS is also part of Eureka ADTT (Advanced Digital Television Technologies). In terms of Philips-CE/BGTV, this report is the FSE (Feasibility Study End) report. As described in the master plan for ADONIS, the FS is concerned with the following subject

Identification of physical parameters and perceptual attributes

which was decided by all ADONIS partners to be subsequently subdivided in the following items:

- Identification of attributes
 - Get hold of descriptive terms, used by potential buyers/naive viewers, in Dutch, in Danish as well as in English
 - Experiments conducted with a wide range of commercially available TV systems
 - Which important dimensions can be recognised within the sensorial image?
- Identification of physical parameters
 - Collection of knowledge about technical possibilities
 - Recognise the important physical parameters
 - Regrouping of technical possibilities into certain classes of physical parameters

At this point Bang & Olufsen is acknowledged for the idea of the general approach of describing the sensorial image, i.e., the descriptive analysis techniques which are well-known in food evaluation research.

In this report the above approach, as it was stated in the master plan, was fully covered. A survey was done w.r.t. known experiments and measuring techniques to collect knowledge about the physical and perceptual world, and theoretical considerations are presented.

Summary

In this report, descriptions of images are given in a fourfold way. The first two alternatives are related to the physical world (light) and to the perceptual world (sensorial attributes), respectively. The second division has to do with natural images versus reproduced images, i.e., in our case outside world images versus displayed images.

Contents

1	Introduction: Framework and Methods	5
2	Natural Images	6
2.1	Physical World	6
2.2	Perceptual World	9
3	Displayed Images	12
3.1	Display Systems	12
3.2	Physical Measurements on Displays	14
3.3	Perceptual Experiments	16
3.3.1	Earlier Investigations	16
3.3.2	Viewing Panel at Philips CE-BGTV, TV Lab Eindhoven	16
	References	23

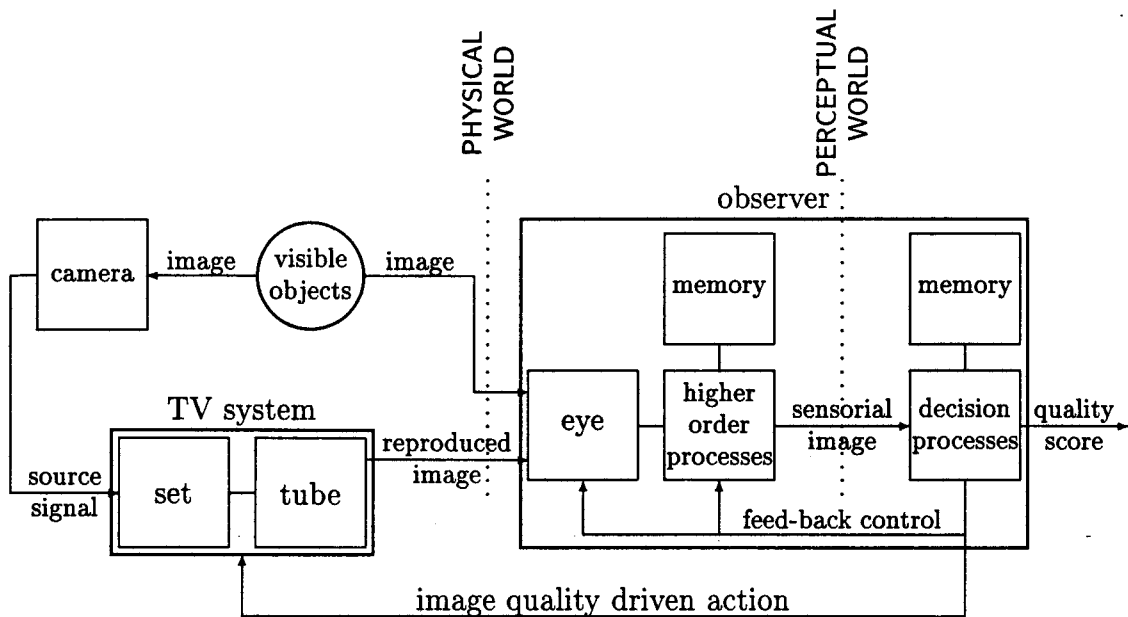


Fig. 1. General model for TV system and observer and the notion of image quality.

1 Introduction: Framework and Methods

Consider the image quality of a TV system. Obviously, the image quality is provided by source signal, set, tube, and viewer, but what is it exactly? Despite the vagueness of the term, it is clear that it must be defined in terms of properties of a 'sensorial image' (see fig. 1). For instance, 100 Hz being better than 25 Hz has to do with the fact that the sensorial image has become stable. Another example: colour images can be quite well reproduced with only three phosphors, which reflects the property of metamerism¹ in the sensorial image.

It becomes clear that if one takes this view, it automatically follows that some representation of the sensorial image should be developed, in terms of language for instance. This requirement is completely analogous to the need of specification of the reproduced image (see fig. 1); in that case too, the right parameters to describe the image might not be as trivial as it might appear at first sight. This report is about both points of view, in terms of measurements and experiments as well as in terms of theoretical considerations. The ultimate problem of quality measurement is postponed to other experiments with larger numbers of subjects.

Nevertheless, the problem with subjective experiments is that the decision block, as indicated in fig. 1, is always present, even if one is not interested in the opinion on image quality but only wants to know the sensorial image (see also [3, 2]). This problem is faced, and the proposed solution is the use of a trained panel, the members of which agree on the notion of a set of terms that describe the sensorial image as accurately as possible. The training also should avoid problems of bias in the process of quantitatively rating the aforementioned subjective terms. How this

¹metamerism = the phenomenon that different spectral distributions of light can lead to the same colour sensation

is solved in practice is still subject to study, which involves questions about absolute/relative reference images/sets and the preferred usage of the evaluation scales (see also section 3.3).

This separation of two tasks, i.e., trying to measure the sensorial image with a trained panel on the one hand, and measure its relation to image quality by consumer tests on the other hand, has also been discussed in [6]. By multi-dimensional methods, notably descriptive analysis techniques which have been decided to be the analysis tools within ADONIS, the two test results can be related. The advantages of considering the additional step inbetween, i.e., the measurement of the sensorial image, are efficiency in optimising image quality, better understanding of human vision, the possibility to systematically create new ideas for improving image quality, and an improved 'resolution' of subjective evaluation of TV systems.

2 Natural Images

Considering the process of seeing, one can start at two sides. First, one can take the mediator of information, i.e., light and consider the first levels of vision from that point. This is what is mostly done. The so-called front-end part of the human visual system is considered as a measuring device of a specific kind, which provides all kinds of (local) measurements of the light distribution that comes in. Second, one can think about what visual perception really is. Obviously, the aforementioned measurements don't mean anything if one doesn't have a kind of internal representation of the outside world (this has been called before the sensorial image, see fig. 1). At this point objects and their properties enter into the discussion: what are these properties, can we make a complete list, and what are the most important ones? This is the perceptual point of view, which is not often the starting point for a discussion.

Both approaches are reflected upon, without any specific reference to how the image comes about. Although the requirements that must be imposed on imaging systems are discussed in the last part of this section, the real restriction to displayed images will not be made until the section on displayed images.

2.1 Physical World

If we see the outside world, it is believed that this seeing is a high level interpretation of the measurement of electromagnetic radiation in a certain window of wavelengths, i.e., *light*. If one thinks about such a measurement with an instrument, it is clear that the intensity of radiation is dependent on the *direction* it comes from, on *time*, on *wavelength*, and on the *position* the measurement is done. The first dependence can be denoted by a solid angle Ω , the second and third by t and λ respectively, whereas the last one is represented by the three-dimensional vector \mathbf{p} . In short, the physical quantity which is subject to measurement when we use our eyes, is the intensity of e.m. radiation:

$$I(\Omega, t, \lambda, \mathbf{p}).$$

Now, every measurement is quantised. Most often, this feature is not noticed, because the level of quantisation is extremely small, but in this case the (non-uniform) quantisation of the dimensions is obvious. Along the dimensions wavelength and position the quantisation is quite dramatic, as the number of sampling points is 3 and 2, respectively, indicating the three different types of cones and the two eyes humans have. As regards direction and time the limits are set by the size of receptors in the retina and the speed of processing of the neurons.

Plain sampling does not describe completely what happens actually in the eye. For a advanced kind of measurement, it would be advantageous not only to know the sample values, but also derivatives thereof; it is observed that something similar takes place in the eye. In fact, a better reconstruction of the original signal can be provided if more orders of the Taylor expansion are known. One can consider this process as a kind of smart oversampling, with different sampling functions, which are derivatives of each other.

The combination of basic sampling functions and their derivatives provides a complete list of all the features that are possibly present in the front-end of the human visual system. Of course, in two dimensions of the intensity function the situation is of limited complexity, because sampling is very coarsely done. For instance, the zeroth order in λ indicates light intensity, whereas first and second order represent the blue-yellow and green-red opponent outputs respectively. For p the first order derivative represents binocular anti-correlation, which leads to the percept lustre. For these dimensions, this listing is complete. However, in higher order, cross-derivatives can be taken, through which the number of basic elements increases quite dramatically.

The previous ideas have been taken from [1] and [8]. The author of the latter article specifically looks into the spatial aspects of vision. He describes sampling and derivatives in the sense of differential geometry, and also considers sampling size as a parameter. Different sampling sizes are hardly ever considered in display-related quality problems, although it is known that specific relations between sampling sizes in different dimensions exist. An example is space and time, where a short time constant is only combined with small spatial scales (regular motion detection) and the coarse spatial scales only are perceptible for longer time constants (long range motion detection system) [9].

The basic sampling functions are different for different dimensions, for instance, the temporal sampling function must obey causality. Nevertheless, apart from the precise form, the set of kernels can be listed if the order of the operators is limited.

A simple list without sampling scale dependence is given in table 1. All separate measurements constitute a complete set if one considers all orders, but they don't have to be in the form the eye preferably uses. For instance, derivatives may be summed or otherwise combined to get alternative measurements. In fact, in table 1 this has already been done in the case of ∂_{u+v} derivatives, which are logical combinations equal to $\partial_u + \partial_v$.

The plenoptic² function approach proposes to only use physical measurements

²*plenus* = complete, *optic* = concerning light

type of derivative	actual dimensions	description	
∂_u	x	vertical 'edge'	
	y	horizontal 'edge'	
	t	flicker (brightening)	
	λ	blue-yellow opponency	
	p_x	binocular anticorrelation ('luster')	
∂_{uv}	xx	vertical 'bar'	
	yy	horizontal 'bar'	
	tt	flicker (pulse)	
	$\lambda\lambda$	green-red opponency	
	xy	criss-cross	
	xt	vertical 'edge' reversing	
	$x\lambda$	vertical 'edge' blue-yellow	
	xp_x	vertical 'edge' anti-correlation	
	yt	horizontal 'edge' reversing	
	$y\lambda$	horizontal 'edge' blue-yellow	
	yp_x	horizontal 'edge' anti-correlation	
	$t\lambda$	blue-yellow reversing	
	tp_x	reversing anti-correlation	
	λp_x	blue-yellow anti-correlation	
	∂_{u+v}	$x + y$	diagonal 'edge'
$x + t$		vertical 'edge' leftward	
$x + \lambda$		vertical hue-sweep	
$x + p_x$		vertical 'edge' with horizontal disparity	
$y + t$		horizontal 'edge' downward	
$y + \lambda$		horizontal hue-sweep	
$y + p_x$		horizontal 'edge' with horizontal disparity	
$t + \lambda$		sequential hue-sweep	
$t + p_x$		sequential eye-order	
$\lambda + p_x$		hue-shift luster	
$\partial_{(u+v)(u+v)}$		$(x + y)(x + y)$	diagonal 'bar'
		$(x + t)(x + t)$	vertical 'bar' leftward
		$(x + \lambda)(x + \lambda)$	vertical hue-sweep
	$(x + p_x)(x + p_x)$	vertical 'bar' with horizontal disparity	
	$(y + t)(y + t)$	horizontal 'bar' downward	
	$(y + \lambda)(y + \lambda)$	horizontal hue-sweep	
	$(y + p_x)(y + p_x)$	horizontal 'bar' with horizontal disparity	
	$(t + \lambda)(t + \lambda)$	sequential hue-sweep	
	$(t + p_x)(t + p_x)$	sequential eye-order	
	$(\lambda + p_x)(\lambda + p_x)$	hue-shift luster	

Table 1. Examples of a periodic table from the plenoptic function approach, as taken from [1]. An overview of basic 'features' for single and double derivatives is given, as well as for tilted first and second order derivatives.

that are performed by early vision too. It would mean for instance, that spectral content is not important anymore, but only the red, green, and blue filtered versions of this measurement.

2.2 Perceptual World

What is the goal of vision, what does it do? As far as I can see it at this moment, the visual system tries to provide information of *objects* in the surrounding world to the organism (see also [5]). Therefore, in order to reflect on perceptual attributes, it is natural to start thinking about properties of objects. A first list can readily be made:

- Objects are *continuously* there
- Objects have a *shape*
- Objects have a *size*
- Objects have *edges*
- Objects can *move*
- Objects have a *depth*, and are *located* somewhere
- Objects have a certain *lightness*, and a certain *colour*
- Objects have a *texture*
- Objects may be *transparent* or *opaque*
- Objects may be *lustrous*
- Objects may provide *direct reflections*

Further, it is important to realise that a certain level of *interest* can alter the requirements which are related to these properties. For instance, one can be urged to come closer, to turn the light on, to follow a movement with the head. The effect of such actions is that the number of subclasses of objects of the object of interest is increased, i.e., the viewer wants to get more information about it.

- Objects have a level of *conspicuousness*

Now we start thinking of the process of reproducing a perceived image. A *display* in itself, when looking at it, is an object which has all the properties mentioned above. However, we do not want to see the display as an object, but reproduce other images with it. Therefore, the conspicuousness of a TV set when in operation, should be very low. It is clear that the *size*, *aspect ratio*, and the *visible frame* around the image are important issues in this respect.

Properties	Requirements	Problems
continuous	brightness stability	flicker noise, snow jitter pumping, breathing
shape, size	geometrical fidelity	raster distortions linearity errors flatness
edges	sharpness, legibility	blur serration, jaggedness curtain (ringing) effect
movement	motion fidelity	jerkiness tailing, motion smear after images shimmer
depth	3d TV	—
lightness	brightness brightness contrast brightness uniformity	dim, grey grey black grey white silting diffuse reflections screen non-uniformity brightness patchiness brightness speckles specular reflections
colour	colour fidelity colour contrast colour uniformity colour sharpness	unnatural complexion white orange red dullness colour patchiness colour speckles, stains colour serration, edges
texture		periodic structures moiré
transparency		
lustre	3d TV	—
reflections		dullness
conspicuousness		
low conspicuousness TV set	size aspect ratio surrounding frame	curved reflections
high conspicuousness image	viewing position independence colourfulness brightfulness	

Table 2. Table of terms related to the 'perceptual image', and reasoning towards the display system.

The reproduction should be 'hifi' in the sense of all properties mentioned above. One can make a list of requirements, whereby the word *fidelity* is added to the terms above:

- Temporal fidelity, i.e., brightness stability
- Shape fidelity, or geometrical fidelity
- Size fidelity (relatively, see before)
- Edge fidelity, which has to do with sharpness and legibility
- Motion fidelity
- Depth fidelity
- Lightness fidelity, which has many aspects like brightness fidelity, brightness uniformity and brightness contrast
- Colour fidelity, and colour uniformity
- Textural fidelity
- Transparency fidelity
- Lustre fidelity
- Fidelity of specular reflections
- Conspicuousness fidelity

It can be argued that the reproduction in itself should have a high conspicuousness, which has to do with high brightness (contrast), colourfulness, and independence w.r.t. the environment. One can compare this to the properties of objects which are conspicuous: they are often nearby, have sharp edges, bright colours, are in the spot light, differ in texture from the environment, move, etc.

Next, object properties, reproduction requirements and common problems are related in a perceptual sense in table 2. This table should be compared to experimentally retrieved listings, such as the one of table 6. The comparison yields a quite good correlation between items, with the exception of the things that aren't present in present day TV systems, such as depth and lustre perception, and some less obvious properties of objects, such as texture, transparency, reflections, and conspicuousness.

The relevant remark that should be made here, is that in the process of linking the descriptive analysis results of the viewing panel to general quality assessments with a large group of consumers, the description of the sensorial image still might not be complete.

3 Displayed Images

3.1 Display Systems

In section 2.1 the stimuli for the visual system were considered as such, and not with a specific imaging application in mind. The changes, or rather limitations, that exist in that case will be discussed in this section.

As is well known, sampling size is crucial for all dimensions: elementary pixel size, temporal refresh rate, bandwidths for colour-difference signals. The sampling which is necessary in the physical world can be read off from the properties of early vision. The most outstanding example for this is the usage of only three phosphors which clearly reflects the three types of cones in the retina. However, the sampling process is not a priori separable across dimensions, and we know this is indeed not the case for the display side as well as for the early vision part. An additional problem is the dependence of sampling size and form on the amplitude of the signal.

For most display types, the first 4 dimensions of the plenoptic function are mapped onto 1 dimension, i.e., time. In that dimension colour, horizontal, vertical, and temporal information are mapped in this order from small to larger time scales³. In order to notice some problems, let us consider a practical example: a tilted red bar on a green background moving to the right is displayed on a screen. The following problems might occur, even if the sampling density in each separate dimension is sufficient:

- a tilted edge cannot be faithfully represented: staircase effect
- the colour difference between the edge and its background competes with the exact position of the edge: anomalous colouring or blurring might be the consequence
- the movement of the bar must be quantised in time and space simultaneously, which cannot be done faithfully: jerky motion or motion blur

Several tricks have been applied to avoid most problems. The best known one is the proposition to sample according to a (3-dimensional) quincunx⁴ lattice in time and space, which takes advantage of the fact that off-dimensional sensitivities are smallest.

What to do with these problems? The most sensible thing is to remember them when considering display properties in terms of the plenoptic function approach. This approaches closely some of the physical measurements that are really done (cf. next section). An example is the measurement of the spot shape, as a function of electron current, which is the basic ingredient to determine how spatially sampled luminance values are interpolated. A simpler example is the measurement of a step response, which reflects the first order spatial derivative in terms of the luminance function.

³In some systems the order is changed. An example is a LC projection system, where colour information is presented 'slower' than spatial information.

⁴*quincunx* = like the arrangement of 5 dots on a die

category	measurements
phosphors	spectral contents temporal behaviour, decay time (x, y) and/or (u, v) measurements white point, blue stretch, etc.
screen	mask (pitch) phosphor lay-out size, aspect ratio viewing distance geometry, curvedness electron scattering glass transmission
signal	frequency characteristics, cut-off frequency off-set synchronisation system (PAL vs. NTSC, etc.) field interpolation, field rate conversion errors (ghost, noise, etc.)
contrast & luminance	MTF measurements large area contrast (dark room; checkerboard) reflection coefficient function ambient lighting conditions luminance (reproduction curve) limiting characteristics uniformity of the screen, moiré, line structure
resolution	luminance profile of spot, dependent on electron current MTF step response # of lines per field, Kell factor # of pixels per line # of fields per second interlace convergence, crosshatch pattern

Table 3. Physical measurements that can be performed on an image as reproduced by a TV system. Not all of them are independent; furthermore, the list might not be complete as it stands.

3.2 Physical Measurements on Displays

The fact that a representation has to be found in physical space is implicitly known to everyone who is active in this field. Possible measurements are listed in table 3, a table which is not considered as being complete, but certainly provides a starting point. A valuable check-list can be found in [13], the ingredients for the SQRI (square root integral) method also belong to this category [7], and the so-called Picture Improvement team within Philips also contributes its part [4].

The counterpart of this list (i.e., for natural images) is presented in section 2.1, which resulted from the plenoptic function approach. Note that some of the measurements are specified in such a way that the step from reproduced (physical) image to sensorial image is as transparent as possible. That is, they mimic the properties of early vision. This is nice, but it also is the source of misunderstanding between technical engineers and people who take the perception point of view. The reason for the confusion is most often that people forget about the step from reproduced image to sensorial image (see figure 1), because they think that all perceptual information is already taken care off by these measurements. As the most outstanding example might serve colour, which has no meaning in the physical world because in the physical world only frequencies and wavelengths exist, but only makes sense in the sensorial image. Only the fact that the eye samples wavelength in a very specific way justifies colour measurements. However, colour perception is not fully specified by only performing colour measurements: colour adaptation and colour constancy are two known phenomena which cannot be explained by just colour measurements. Another example is luminance: luminance measurement is already an intensity measurement, which is non-uniformly weighted over wavelength, in order to mimic the eye sensitivity. However, luminance is only one factor in understanding brightness perception, i.e., the knowledge about why people experience one object brighter than the other.

Summarising this, one can state that many measurements try to involve already the first part of the observer, a part which is physiologically best known and specified (e.g., the CIE colour spaces).

The use of so-called Artificial Intelligence in TV systems can make the representation of the physical 'world' extremely difficult, especially if one doesn't know the exact functionality of a specific AI implementation under study. However, in the latter case general knowledge cannot be built up anyhow. At this point I would like to formulate a premise which is always implicitly understood in this project: it must be possible to make an explicit physical representation of the system under study, we do not want to study black boxes. If this premise is not fulfilled (e.g., if one has a set with a neural network which is not specified), one can still evaluate the system in a better way (with a viewers panel), but the relation between physical and perceptual cannot be found.

category	citations
sharp image	sharp, also in the corners not unnatural sharp, but nature-like, natural reproduction is optimal image without any deformations, circles are circles, not eggs sharpness is related to colouring, reproduction of bright and darkness images shouldn't swim no twin images, no distortions
colour	natural, nature-like colours, natural colour diversity clear, strong, saturated colours, not mixed, but conspicuous no colour forgeries, but real colours sharp, non-flickering colour edges, clear segmentation neutral and exact colours unique colour contrasts without grey haze in colours without bluishness colours in harmony with each other, 'round' combinations eye-friendly colours, comfort individual colour tuning
restful image	without flicker, red is extra sensitive higher refresh rate sharp images, without hurting the eyes without noise, clear image free of snow, sharp cut image visible brightness variations, image gives no rest good static image Sony Trinitron System, better for the eye
contrast	black-and-white contrast, discrimination between brightness steps contrast sharpness, high contrast, good contrast resolution contrast: more pixels, better image sharp and clear contours, foreground and background separation clear contours, even for the smallest details, detail sharpness
screen	free of reflections, antistatic for cleanness larger, flat screens for exact detail reproduction non-rounded corners for shape fidelity
reception	cable transmission for clear image antenna independent, distortion free reception clear image reception, for far transmissions too

Table 4. Results from the SINUS investigation [11]. These are citations of subjects when asked after associations with the term 'image quality'. The categorisation was done (afterwards) by the experimenter.

3.3 Perceptual Experiments

3.3.1 Earlier Investigations

ADONIS is not the first project in which the perceptual nature of image quality is considered most important. Within Philips, I think of two earlier investigations, i.e., the 'SINUS' customer inquiry [11] and the 'Starry Skies exercise' [12]. Especially the first one contains a lot information about how to describe the sensorial image. In table 4 the terms are cited which were found by just gathering citations of many subjects on image quality.

The second investigation was a workshop of several days on the whole problem of ADONIS and a little more. The recommendations which this team put forward contain amongst others the institutionalisation of a common language for picture quality, and the training and use of test panels for subjective testing, exactly what we are trying to do in ADONIS at this moment. The results of the workshop are a list of subjective terms, sometimes mixed with physical terms, a rating of importance of the items on that list, and schematic relations between the most important subjective terms and their physical counterparts. The first list mentioned is presented in table 5.

The point of view of the Starry Skies team is somewhat more on the physical side, but nevertheless they find important additional factors of (image) quality which is mainly due to the fact that marketing people were involved.

3.3.2 Viewing Panel at Philips CE-BGTV, TV Lab Eindhoven

In ADONIS we had two viewing panels set up, one at Bang & Olufsen and the other at Philips Consumer Electronics. The reason for a viewing panel is the following: if one wants to know the relation between a physically specified image (display side) and the perception thereof (sensorial image), the found relation can only be as accurate as the least well known side. As the existing measurements on displays are quite elaborate, it follows that the sensorial image has to be measured in an as accurate as possible manner. The only way to reach this goal is to use a trained panel (see also [6]). The reasoning of having 2 panels instead of 1, is practical as well as fundamental. First, both B&O and Philips want to evaluate own collections of TV systems. Second, the panels develop descriptive terms, which are language specific, and we want to have an additional check within ADONIS against possible erroneous results.

At this moment, the results of the panel at Bang & Olufsen are not ready for publication in this report.

Up to now, the procedure of setting up a panel at Philips CE-BGTV, in the TV lab Eindhoven, can be summarised as follows:

- Panel members were selected from non-technical people within the TV lab. At this moment the panel consists of 9 members, 5 female and 4 male. The women in the panel are all secretaries, whereas the men are archivist, lay-outer, purchaser, and manager of a TV test room. Up till now, the panel members were not yet tested on normal vision.

category	subdivision
natural colours	white colour point red colour point green colour point blue colour point colour saturation colour deviation in position
flicker	large area flicker movement of details cross colour cross luminance
flatness	horizontal radius vertical radius
format	aspect ratio screen size styling
picture noise	signal-to-noise in luminance noise in colour
sharpness	pseudo 3d-effect
readability	character recognition test level of fatigue related to TV viewing time availability of correct lettersets
brightness	peak brightness average brightness
contrast ratio	
external conditions	daylight viewing range external light intensity colour distribution of external light
TV set	depth of TV difference in colour tone between screen & total design colour tone of screen
miscellaneous	independence perceived picture quality from content effective screen size related to visible screen size on-time independence perceived picture quality from audio quality

Table 5. One of the Starry Skies workshop results, which indicates the terms this team associated with picture quality [12]. The categories as well as their subdivisions are developed within the workshop.

Dutch term	Dutch description	English term	English description
kleur-hoeveelheid	de mate van kleur in een plaatje	colourfulness	the amount of colour in an image
kleur-echtheid	natuurgetrouw, eerlijk, afwezigheid van kleurvlekken (kleurvlakken/vlakjes die niet in het beeld thuishoren)	colour genuineness	natural, honest, without any colour stains (colour areas that don't belong in the image)
kleur-plaatsing	scherpe overgang tussen aangrenzende kleuren (kleuren mogen niet in elkaar overlappen), de mate van matching van de kleuren op de objecten (kleur mag niet half naast het object liggen)	colour positioning	sharp transition between adjacent colours (colours shouldn't blend), measure of matching of colours with the objects (colour shouldn't be shifted w.r.t. the object)
lichtheid	hoeveelheid licht van de beeldbuis ten opzichte van de omgeving	lightness	amount of light of tube w.r.t. the environment
donkerheid	de mate van donkerheid van de donkere delen in het beeld (de donkere delen kunnen te licht of te donker zijn)	darkness	measure of darkness of dark parts in the image (they can be too bright or too dark)
lichtdetail	waarneembaarheid van details in de lichte delen van het beeld	light-detail	perceptibility of details in the bright parts of the image
donkerdetail	waarneembaarheid van details in de donkere delen van het beeld	dark-detail	perceptibility of details in the dark parts of the image
scherpte	stijlheid van de donker/licht overgangen (denk aan een scheermes of een haarscherpe lijn), leesbaarheid van een tekst	sharpness	steepness of dark/light transitions (think of a razor-blade or dead sharp line), legibility of text
beeldrust	afwezigheid van: strepen, trillingen, storingen, beeldflakker (flikker), kleurkriebels in geruite jasjes, ruis	restful image	absence of: stripes, vibrations, distortions, flicker, itching colours in chequered suits, noise
bewegingsweergave	goede weergave van bewegingen zonder rare neveneffecten (bv. 2 voetballen i.p.v. 1, of een schokkende beweging i.p.v. een vloeiende beweging)	motion rendering	rendering of motion without peculiar effects (such as 2 soccer balls instead of 1, or jerky motion instead of fluent motion)
vormechtheid	rechte lijnen (bv. deuren) moeten recht zijn en niet krom	shape genuineness	straight lines (e.g. doors) should be straight and not curved

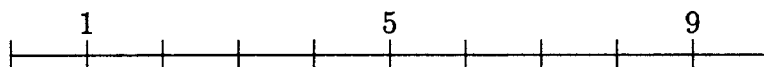
Table 6. First list of factors that constitute image quality, as developed in the viewers team of Philips CE-BGTV TV lab Eindhoven. The original Dutch terms are given, with Dutch descriptions, and their closest English translations.

- The panel members were asked to write down individually and without watching TV sets all factors that determine image quality in their opinion.
- The results were put together and discussed in a plenary session. In this session (still without TV sets) it was decided which terms were the same, and, when the factors were differently described, a common term was chosen. This stage took two sessions of one-and-a-half hours each. The result of this process is listed in table 6.
- The panel evaluated the descriptive terms on the following TV systems, which were playing with fixed chosen factory pre-settings (AI 'high', picture quality 'normal', noise 'off', digital scan 'off', line flicker red 'on', sharpness 'neutral'):

nr.	brand	type	f (Hz)	tube
1	Sony	KV-S2941D	100	Super Trinitron (SVM)
2	Grundig	M72-105IDTV	100	Toshiba (DAF,SVM)
3	Philips	29PT910B	100	BLS (SVM)
4	Panasonic	TX29A3C	50	GAOO A3
5	Telefunken	S580	50	Black DIVA
6	B&O	MX7000	50	BL
7	Loewe	Arcada 72	50	
8	Philips	28PT530B/00	50	BLS

From the analysis of the results it appeared that some panelists had evaluated the strength of the perceptual attributes (that's what we wanted them to do), but others had evaluated the influence of the specific attributes on quality, what is unfortunate, because in that case one loses information on the 'step in between', i.e., the sensorial image (cf. fig. 1). Therefore, after a plenary discussion with the panel, where this problem was discussed in length, the experiment was repeated.

- The evaluation was done on vertical graphical scales of 10 cm long, 8 scales per sheet, one factor per sheet. The 8 lines were used for the 8 different TV systems. An example of such a line is given below, albeit horizontally displayed.



The use of adjectives is avoided because it could have introduced bias in the scores [10].

Analysis of the results of the above mentioned experiment shows that the panelists still have several problems with the task of graphically scaling the attributes.

First, the usage of the scale is different for different subjects. The way to correct for this is so-called z -transformation. The rationale behind this transformation is

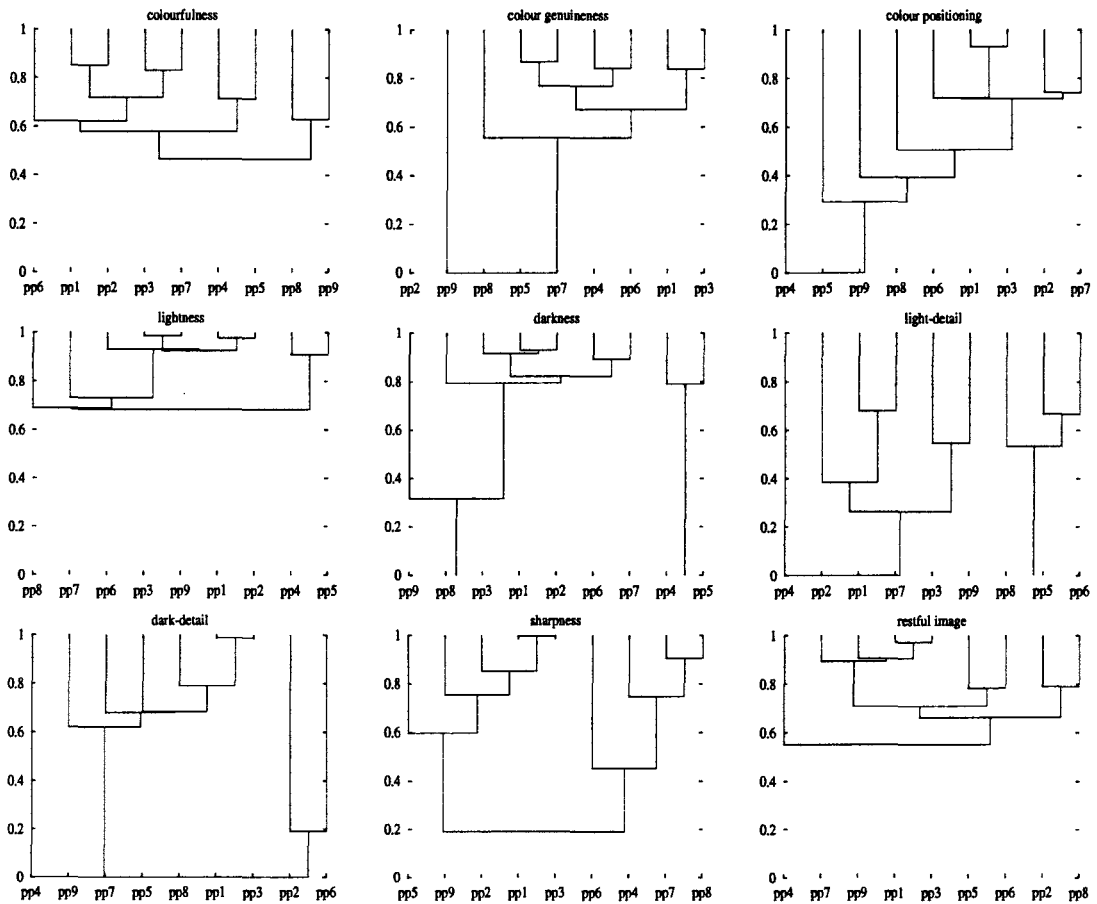


Fig. 2. Clustering on basis of correlations between subjects. This is done for each attribute separately.

that the differences in usage of scale are caused by the decision block in fig. 1. The z-transformation linearly transforms the scales in such a way that the means and the standard deviations of the individual subjects (over all their observations) become equal to each other.

Second, the individual notions of the attributes are not in line with each other. One can analyse this by calculating the correlations between subjects, whereby the scores of one attribute for all TV sets are compared. Leaving out the last two attributes, motion rendering and shape genuineness, where the majority of panel members didn't score at all, the correlations were quite low. For 4 of the 9 remaining attributes there was at least one panel member who did not correlate at all with all the others. The so-called clustering of panelists on the basis of correlations can be found in figure 2, where the last two attributes were left out. Note that all attributes suffer from a non-unique notion between subjects. Further note that these correlations are not sensitive to any linear transformation, i.e., linear variations in usage of

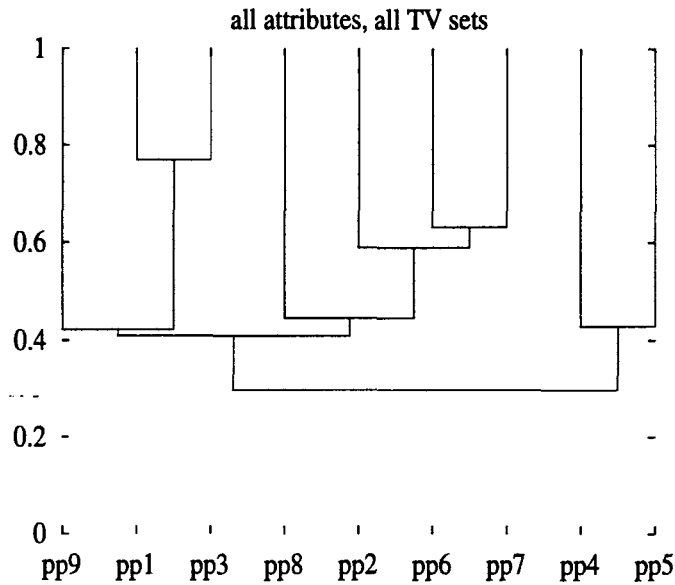


Fig. 3. Clustering on basis of correlations between subjects over all scores they gave.

scale don't matter within this analysis.

Third, one can repeat the previous argument if one does the same analysis over all attributes and TV sets together. It means, that if one correlates the 88 scores per subject with each other, several subgroups in our panel appear. In our analysis (see figure 3), panelists 1 & 3, and 2, 6 & 7 are two subgroups, whereas the others appear individually. The conclusion is that a lot of training on scale usage and attribute alignment has to be done.

Fourth, one can assume that even if the panelists didn't use the scale as well as we'd like them to do, they should have ordered the sets all right on the scale. And maybe these orderings are the same for some attributes. In short, one can perform a factor analysis on qualitative data, which is known as MDPREF analysis (Multi Dimensional Preference). One can do separate analyses per subject and per attribute, but only one analysis will be presented here. We z-transformed the scores, averaged them over the subjects, and used the orderings of the averaged scores to do our analysis, the results of which are plotted in figure 4. The interpretation of this plot is as follows. The attributes were linearly recombined in 11 alternative factors in such a way that the first combination could explain most variation in the data, the second combination could explain the largest part of the remaining variation, etc. It appeared that in this case 2 factors were sufficient to explain 98% of all variation in the ordering data. The positions of the attributes in this plot indicate how they contribute to these factors. For instance, colour genuineness is almost exclusively described by the first factor, whereas shape genuineness is included in the first as well as in the second factor. Further, the scores of the TV sets can be determined

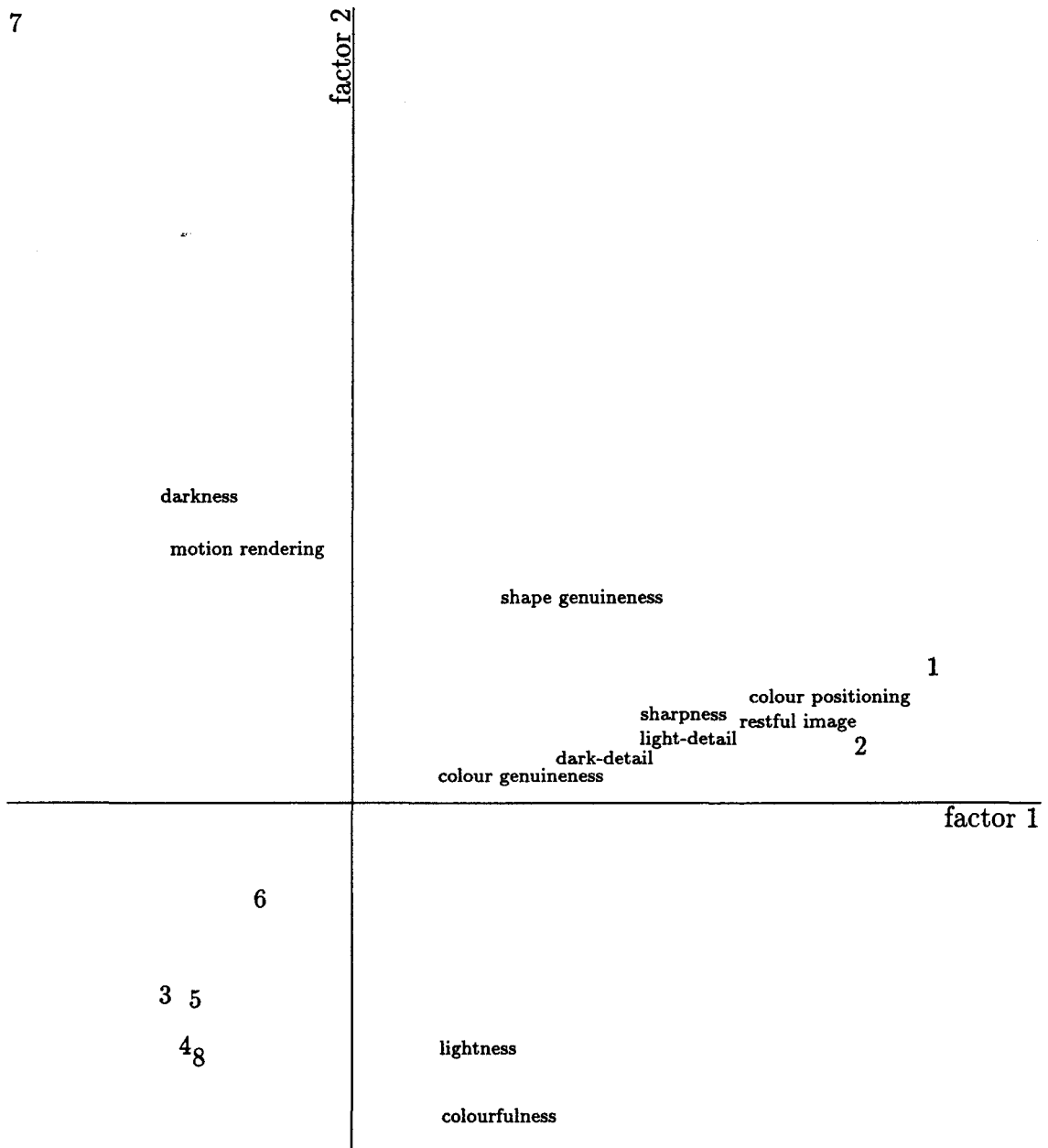


Fig. 4. MDPREF plot of attributes determining directions and TVs, derived from z-averaged scores over all subjects.

for these 2 factors. They are also given in this plot. The combination of attribute position and the positions of the TV sets gives the ordering of TV sets for that attribute. One can do this by imagining a line from the origin of the plot through the position of an attribute, and then projecting the positions of the TV sets onto that line in a perpendicular way. The resulting ordering on that line represents the ordering of TV sets for that particular attribute.

Fifth, one can analyse the results by an analysis of variance on each attribute, whereby the proposed model is such that variations are explained by either TV sets or panel members, but no interaction is assumed. The results then show that panelists' variance is larger than the one of the TV sets for 6 out of 11 attributes. Only in the cases 'lightness' and 'restful image', the TV sets constitute the 'only' and most important effect. The conclusion has been stated before: a lot of training is necessary.

References

- [1] E.H. Adelson and J.R. Bergen. The plenoptic function and the elements of early vision. In M. Landy and J. Morshon, editors, *Computational models of visual processing*, page 3. MIT Press, 1991.
- [2] F.G. Ashby. *Multidimensional models of perception and cognition*, chapter 16, pages 449–483. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1992.
- [3] F.G. Ashby and J.T. Townsend. Varieties of perceptual independence. *Psychological Review*, 93:154–179, 1986.
- [4] P.J. Barth. Picture Improvement, an overview of the fundamental system aspects. Technical report, TV lab Eindhoven, Philips CE, September 1994.
- [5] I. Biederman. Recognition-by-Components: A Theory of Human Image Understanding. *Psychological Review*, 94:115–147, 1987.
- [6] R. Hamberg. Food evaluation: new methods for image perception research? California, USA, June 1994. *IPO*, Reisverslag no. 205, 1994.
- [7] P. Houben. SQRI calculations, sharpness measurements and mini poll results 25" TTV sets. Technical Report ADONIS document 15, ITC, Philips Components, August 1994.
- [8] J.J. Koenderink. What is a "Feature"? *J. of Intelligent Systems*, 3:50–81, 1993.
- [9] K. Nakayama. Biological image motion processing: a review. *Vision Research*, 25:625–660, 1985.
- [10] E.C. Poulton. *Bias in Quantifying Judgments*. Lawrence Erlbaum Associates, Hove and London, UK, 1st edition, 1989.
- [11] SINUS. Erlebnis von Bildschirmqualität, Marktpsychologische Untersuchung. Technical report, Forschungsgesellschaft für Umwelt, Technik und Arbeit mbH, Eindhoven, November 1990.
- [12] J. Swinkels, N. Hermkens, and R. Gal. Starry Skies – Exercise, Picture Quality BGTV. Technical report, CFT Development Support, Philips, 1994.
- [13] J.H.D.M. Westerink, T. Doyle, and F.H. Fleuren. A checklist of perceptually relevant parameters for the characterization of displays. *IPO*, R721, 1989.