Perceptual assessment of sampled images

van Boxtel, P.A.M.

Award date:
1993

Eindhoven University of Technology

MASTER

Perceptual assessment of sampled images

van Boxtel, P.A.M.

Award date:
1993

Institute for Perception Research
PO Box 513, 5600 MB Eindhoven

19.11.1993

7080

Rapport no. 946

Perceptual assessment of sampled images

P.A.M. van Boxtel

Afstudeerder : P.A.M. van Boxtel
Begeleider : Dr.ir. J.B. Martens
Afstudeerhoogleraren : Prof.dr.ir. J.A.J. Roufs
Prof.dr.ir. P. Eykhoff

Summary

If we want to reproduce an image on a display, we have to sample it. The samples have to be chosen such that the perceptual quality of the reproduced image is maximized. The perceived image is modeled as an interpolation of the sampled image. The interpolation function is determined by the fixed display response and the, viewing distance dependent, impulse response of the visual system. It is represented by a B-spline function. For the interpolation function an optimal sampling function can be determined, according to a vector space model. If the image is filtered with this sampling function prior to sampling, the mean squared error between the original image and the reproduction is minimized.

However, this error criterion is not a good measure of perceptual quality. To determine the perceptual quality, psychophysical experiments are done. These experiments have a twofold purpose. Firstly, they have to give insight in the way an image is impaired by sampling and interpolation. Therefore a multidimensional scaling experiment is done. From the scaling results of this experiment, a psychological space is constructed. This space shows how the perceptual attributes ringing and sharpness influence the image quality. As a result follows that the image quality is mainly determined by the sharpness of the image.

Secondly, different sampling functions are compared to find the function that gives the maximum quality. However, it was not possible to compare the original images, which are on the slides, with the reproductions, which are on the display, under equal viewing conditions. Therefore is also determined how the quality judgment is affected when subject are not able to compare the reproduction with a reference image. It appeared that when there is no reference image, maximum quality is reached with a reproduction that is sharper than the reference. The perceptually optimal sampling function is a function that increases both the strength of ringing in the image and the sharpness of the image.

If an image has to be shown on a display, best results are obtained if it is filtered with a first order B-spline sampling kernel, prior to sampling. However, the gain of doing so is not very high, because the scanner has very low noise.

Contents

Chapter 1

Introduction

Suppose we have an photographic image on a black and white slide and we want to reproduce that image on an electronic display. Our aim is to optimize the quality of the reproduction. But how should we define the concept of quality? We can use a mathematical definition like the mean squared error between the luminance pattern on the display and the light transmittance of the slide. However, a human observer does not perceive quality according to such a criterion. When looking at the display he will get an impression of an image. This impression he will relate to a reference. This can be his daily experience, but it also can be a physical reference, like the original image on the slide. We will define image *quality* by comparing the perceived image with a physical reference. When we speak of the *preference* of the perceived image in this report, there is no such physical reference.

The image on the display is discrete, it is built up of a finite number of display elements, called dels, which are usually positioned on a rectangular lattice. The dels have fixed luminance patterns, with variable amplitudes. These amplitudes are controlled by the grey-values, or pixels. The result of the individual luminance patterns is the image one sees. The perceived image is modeled as an interpolation of the pixels. The pixels have to be chosen such that the quality of the image is optimized. They are obtained by sampling the image of the slide. Sampling converts the original photographic image into a discrete image. How to find the pixels values is the central problem of the research discussed in this report.

The sampling and interpolation process can be formulated mathematically. Sampling is performed by filtering the photographic image and measuring the result at discrete points. The impulse response of the filter is called the sampling function. The second part, measuring the filter output at discrete points, is referred to as down-sampling. The samples are used to reconstruct the image as good as possible. This is done by an interpolator, that multiplies each sample with an interpolation function and adds the results. In figure 1.1 sampling and interpolation of a one-dimensional signal are depicted schematically.

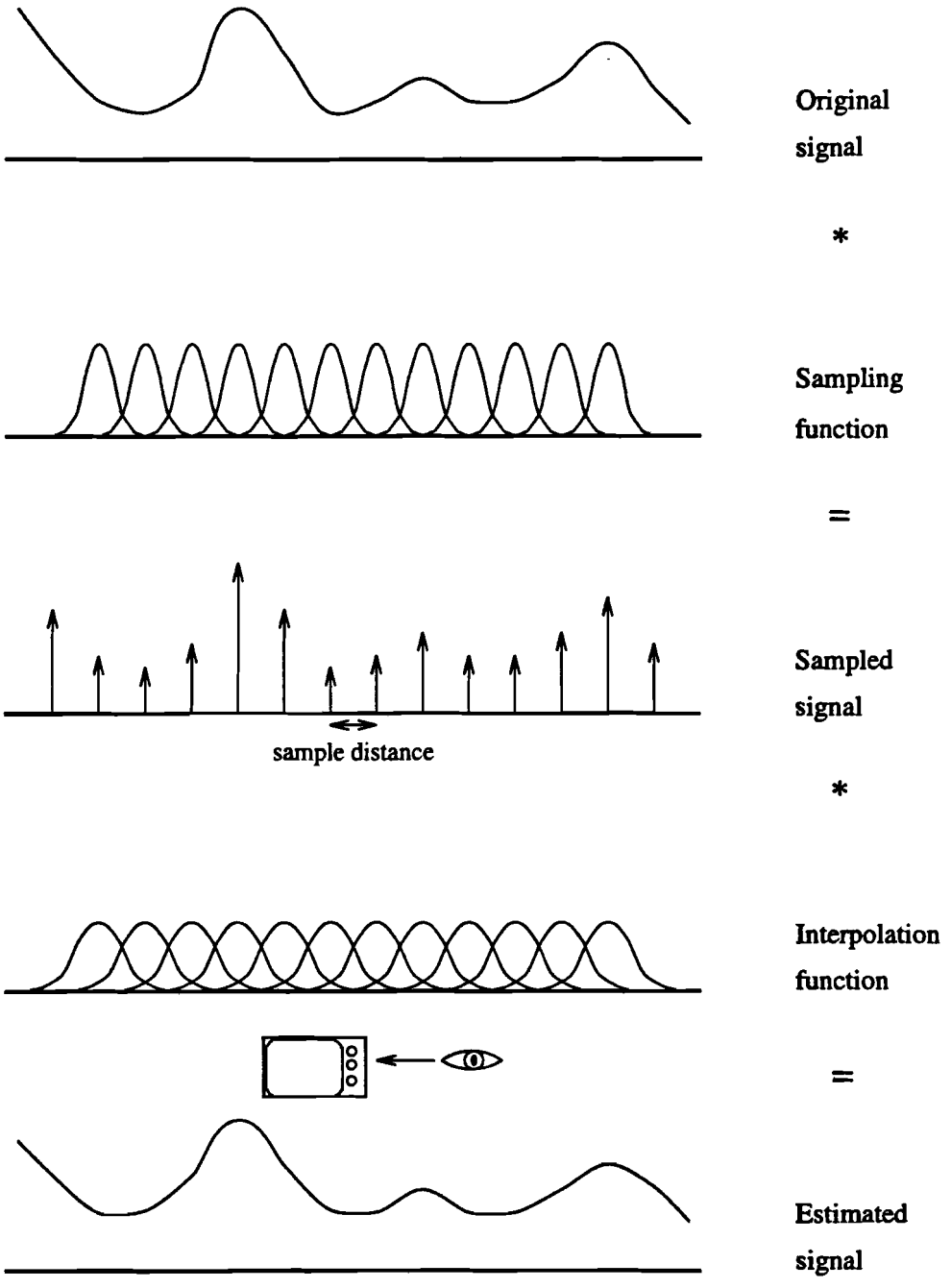


Figure 1.1: Sampling and interpolation of a one-dimensional signal

It is stated by the Shannon sampling theorem that a bandlimited signal can be perfectly reconstructed from its sample values if a $\sin(\pi x/\Delta)/(\pi x/\Delta)$ function is used for both sampling and interpolation [Shannon 1949]. If the signal is not bandlimited then the sample distance Δ determines the accuracy of the reconstruction. Using the theory of lattices this theorem can be generalized to a sampling theorem for N-dimensional signals [Peterson 1962, Dubois 1985]. However, in a display situation the interpolation function is given. It is of finite extent and differs significantly from a $\sin(\pi x/\Delta)/(\pi x/\Delta)$ function. Therefore we have to let go classical sampling theory.

Instead we try to answer the following question. Suppose that the interpolation function is fixed and given. How should we choose the sampling function in order to minimize the mean squared error between the original and the reconstructed image? Once we have chosen a sampling function this way, we have to determine whether it is perceptually optimal, since the mean squared error is not an accurate model of the perceived impairments in an image. Through psychophysical experiments the quality of images can be determined. However, when we show an image on a display, comparison with the original image on the slide is impossible, because we are seldom able to show them both under equal viewing conditions. To be able to show a reference image, we simulate the viewing situation. This means that we sample ánd interpolate the image and take the unprocessed image as a reference. Now we can measure both preference and quality and we can find out how they are related.

But first we want to know what actually happens to an image when we sample it. What is the influence of the artefacts, introduced by sampling and interpolation, on the image quality? It is known that sampling and interpolation gives rise to four artefacts; aliasing, visibility of sampling structure, blur and Gibb's phenomenon, also known as ringing [Reitmeier 1981]. As we will see, most sampling functions have low-pass filtering characteristics, so that the effect of aliasing will be limited. In practice visibility of sampling structure can be reduced by increasing the viewing distance up to the point where the individual dels merge. Increasing the viewing distance will increase the relative impact of the neuronal and optical processing, due to the smaller distance between the dels. Therefore the perceived image will be blurred. The task of the sampling function can be interpreted as minimizing this blurring for a given viewing distance. Thus the optimal sampling function depends on the viewing distance. The disadvantage of the theoretical optimal sampling function is that it usually introduces ringing. By means of multidimensional scaling the perceptual influence of the artefacts ringing and unsharpness on the image quality can be determined. Multidimensional scaling is a method for comparing stimuli that vary in more than one aspect.

This report is organized as follows. In the second chapter we model the display situation. The luminance pattern of the dels and the visual system response are modeled by one interpolation function. B-spline functions are introduced as approximations to realistic interpolation functions. To find the optimal sampling

function, a vector space model for image sampling and interpolation is adopted. In chapter three this model is described and applied to the specific case of B-spline interpolation. In chapter four the multidimensional scaling experiment is described. In chapter five different sampling functions are compared using simulated interpolation functions. This way also the relation between the quality judgment of an image and the preference is examined. In chapter six different sampling functions are compared in a more realistic viewing situation. In the last chapter some conclusions are drawn from this research and some suggestions for future research are made.

Chapter 2

The display situation

We start with a front-to-end analysis of the display situation. This situation is modeled in figure 2.1. First the image of the slide is densely sampled to a discrete image with a very small sample distance. This is done by means of a high resolution slide scanner. The densely sampled image is sampled to a coarsely sampled image, using a sampling filter and a down-sampler. The sampling filter has to be optimized, as will be discussed in the next chapter. Starting from discrete images obviously gives us better control of the sampling process. Finally the sampled image is shown on a display.

The display situation can be modeled by interpolating the discrete image. In the figure this is done by an up-sampler and an interpolation function. The up-sampler pads the discrete image with zero values, to produce a continuous image. The interpolation function is made up of two parts. Firstly, the display function interpolates the sample values to dels. Secondly, the visual system operates on the displayed image. We will model both components of the interpolation function as gaussian functions, so that the interpolation function is also a gaussian function. Analoguous to the in statistics well-known central limit theorem, it can be assumed that the interpolation function will be approximately gaussian, even if both components are not perfectly gaussian. Because the visual system is included in the interpolation process, the interpolation function depends on the viewing distance.

In this chapter, the display situation is analyzed. In the first section, the scanning process is described. In section two the interpolation function is modeled as a function of the display function and the impulse response of the visual system. Both the relation between the densities at the slide and the output of the scanner and the relation between the pixels and the amplitude of the dels are not linear. In section three is demonstrated how this affects the interpolation function. In the last section B-splines are introduced. They are used to approximate the interpolation function.

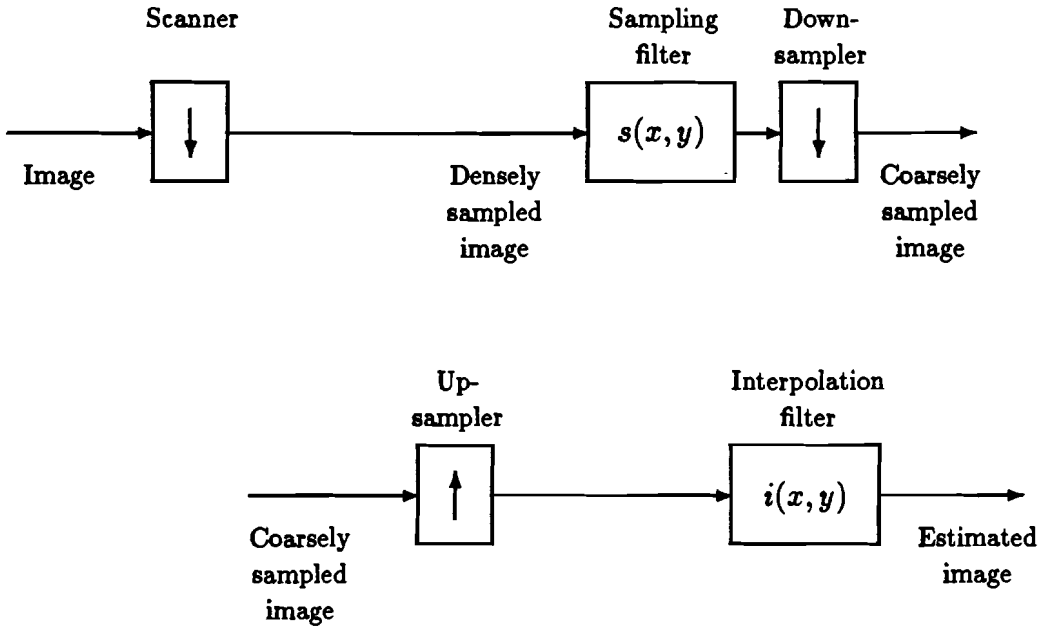


Figure 2.1: The complete model of the display situation

2.1 The scanner

The high-resolution scanner[1] measures the light transmittance of the slide at rectangularly positioned points. It has a sampling function that can be approximated by a delta function. The local densities are linearly converted to 16-bit discrete numbers. These numbers are converted to 8-bit grey-values, so that the resulting discrete image is also quantized. However, the error introduced by this quantization is usually not perceivable. Therefore we will ignore this quantization in the rest.

The characteristics of the conversion from 16-bit values to 8-bit grey-values can be adjusted. It is possible to perform a so-called gamma correction. In that case the conversion is not linear. In a later section this non-linearity will be discussed. We have given gamma a value of one. The densities are scaled linearly between 0 and 255. The lightest and darkest point on the slide are assigned the grey-values of 255 and 0 respectively.

The slides have sizes of 36 by 24mm. From this slide a rectangular area of 24 by 24mm is used to produce a densely sampled image with 2048 by 2048 pixels. Now the sampling distance is approximately 12μm.

[1] Leafscan-35. Leaf Systems Inc., Southboro, Massachusetts, USA.

2.2 The overall interpolation function

The overall interpolation filter is the cascade of two filters with gaussian impulse response. The first is the fixed display function $g_d(r)$. The second is the impulse response of the visual system, which is a function of the viewing angle. We will express the interpolation function as a function of the viewing angle.

On the display each grey-value is mapped to one del. A del has a fixed luminance pattern, with an amplitude that is determined by the grey-value. The relation between amplitude and grey-value is usually non-linear, as will be discussed later. More important at this moment is the shape of the del's luminance pattern. For the reason mentioned above it would be nice if we can model it by a gaussian function, i.e.

$$g_d(r) \propto e^{r^2/\sigma_d^2}, \tag{2.1}$$

in which σ_d is much smaller than half the distance between two dels. This function is the display function, which is the impulse response of the display.

The impulse response of the visual system can be represented by the foveal point spread function, which is a sombrero like function [Blommaert 1981]. However, we will represent it by the eye's optical point-spread function. This function is modeled as a gaussian function with a spread of 0.5 min. of arc. The gaussian function is not a completely adequate model, because theoretically it has infinite extend, but this will be dealt with in the last section.

Modeling the impulse responses by gaussian functions, is obviously a pragmatic choice. The convolution of the two gaussian functions is

$$g(r) = g_d(r) * g_v(r) = e^{-(r/\sigma_d)^2} * e^{-(r/\sigma_v)^2} = e^{-(r/\sigma)^2} \tag{2.2}$$

with

$$\sigma^2 = \sigma_d^2 + \sigma_v^2, \tag{2.3}$$

which is a gaussian function again. Looking at each function separately, this model may not be very accurate. However, if we cascade several filters with arbitrary impulse responses, then the result will converge to a gaussian. So why not make life easy and model each filter separately by a gaussian function?

We will illustrate the models presented above for a standard viewing situation. Suppose the image shown has a size of 512 by 512 pixels and the viewing distance is six times the image height. In this case, the sample distance is $\delta = \arctan(1/(6 \cdot 512)) = 1.1$ min. of arc. It is known that for such large viewing distance, the luminance pattern of the dels is invisible and that the display function can be neglected.

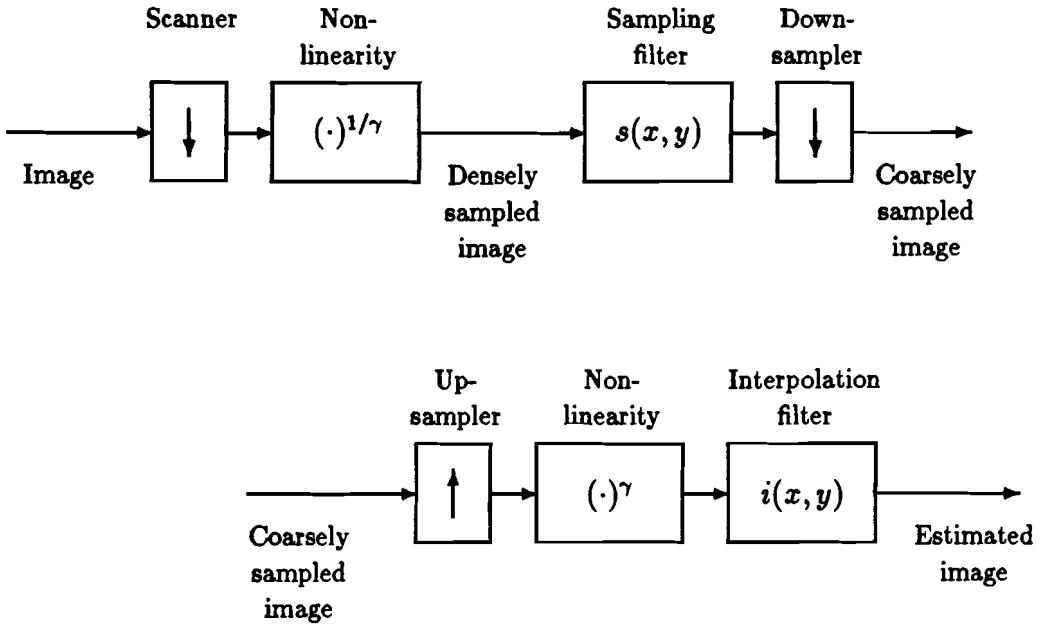


Figure 2.2: The complete model of the display situation, including the non-linearities

2.3 Non-linearities

If we show a reproduction of a slide on a display, we want the luminance of the reproduction to be linearly related to the transmittance of the slide. Generally a display does not have a linear characteristic. The pixel $f_s(x,y)$ and the peak luminance $L(x,y)$ of the del are approximately related by the power law

$$L(x,y) \propto (f_s(x,y))^{\gamma_d}, \tag{2.4}$$

in which γ_d typically has a value of 2.5. To compensate for this effect, the densities of the slide $D(x,y)$ and the output of the scanner $f(x,y)$ are related according to the power law

$$f(x,y) \propto (t(x,y))^{\gamma_s}, \tag{2.5}$$

with $\gamma_s \approx 1/\gamma_d = 0.4$. So the net gamma is one. Furthermore, perceptual tests have shown that the preference of the reproduction is higher when the value of gamma is greater then one.

These power law relations introduce some non-linearities in the model of the display situation in figure 2.1. Prior to the sampling filter there is a compression operation according to (2.5) and prior to the interpolation filter there is a decompression operation according to (2.4). This results in the situation drawn

in figure 2.2. If we are able to interchange the decompression operator and the interpolation filter, then the cascade of the sampling and interpolation operators is linear.

When we interpolate a sampled signal $f_s(n)$ with a gaussian interpolation function $g(x)$, the result is the real signal

$$\tilde{f}(x) = \sum_n f_s(n) \cdot g(x-n). \tag{2.6}$$

We now use the approximation

$$\left(\sum_n f_s(n) \cdot g(x-n)\right)^\gamma \approx \sum_n f_s^\gamma(n) \cdot g^\gamma(x-n). \tag{2.7}$$

This approximation holds when the spread of the gaussian interpolation function is such that the overlap of the two gaussians functions, that interpolate neighbouring samples, is negligible. Gaussian functions have the property that their shape is not affected by a power transformation, i.e. the result of a power transformation with exponent γ_d on a gaussian function with spread $\sigma\sqrt{\gamma_d}$ is a gaussian function with spread σ. So with this approximation the decompression operator and the interpolation filter in figure 2.2 can be interchanged if the gaussian interpolation function with spread σ is replaced by a gaussian function with $\sigma\sqrt{\gamma}$.

2.4 B-splines

Gaussian functions have the disadvantage that they have infinite extent, which makes them hard to implement on a computer. Therefore we try to find another class of functions with gaussian-like shapes and properties. These demands are satisfied by the B-splines, introduced by [Schoenberg 1946, 1969]. B-splines are bell-shaped piecewise polynomials, which originate from multiply convolving a sample-hold function with itself. The B-spline of order n is defined by

$$\beta^n(x - \tfrac{1}{2}(n+1)) = \sum_{j=0}^{n+1} \frac{(-1)^j}{n!} \binom{n+1}{j} (x-j)^n \mu(x-j), \tag{2.8}$$

where $\mu(x)$ is the unit step function

$$\mu(x) = \begin{cases} 1 & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}.$$

Like gaussian functions, these B-spline functions satisfy the convolution property

$$\beta^n(x) = \beta^{n-1}(x) * \beta^0(x) = \beta^0(x) * \cdots * \beta^0(x). \tag{2.9}$$

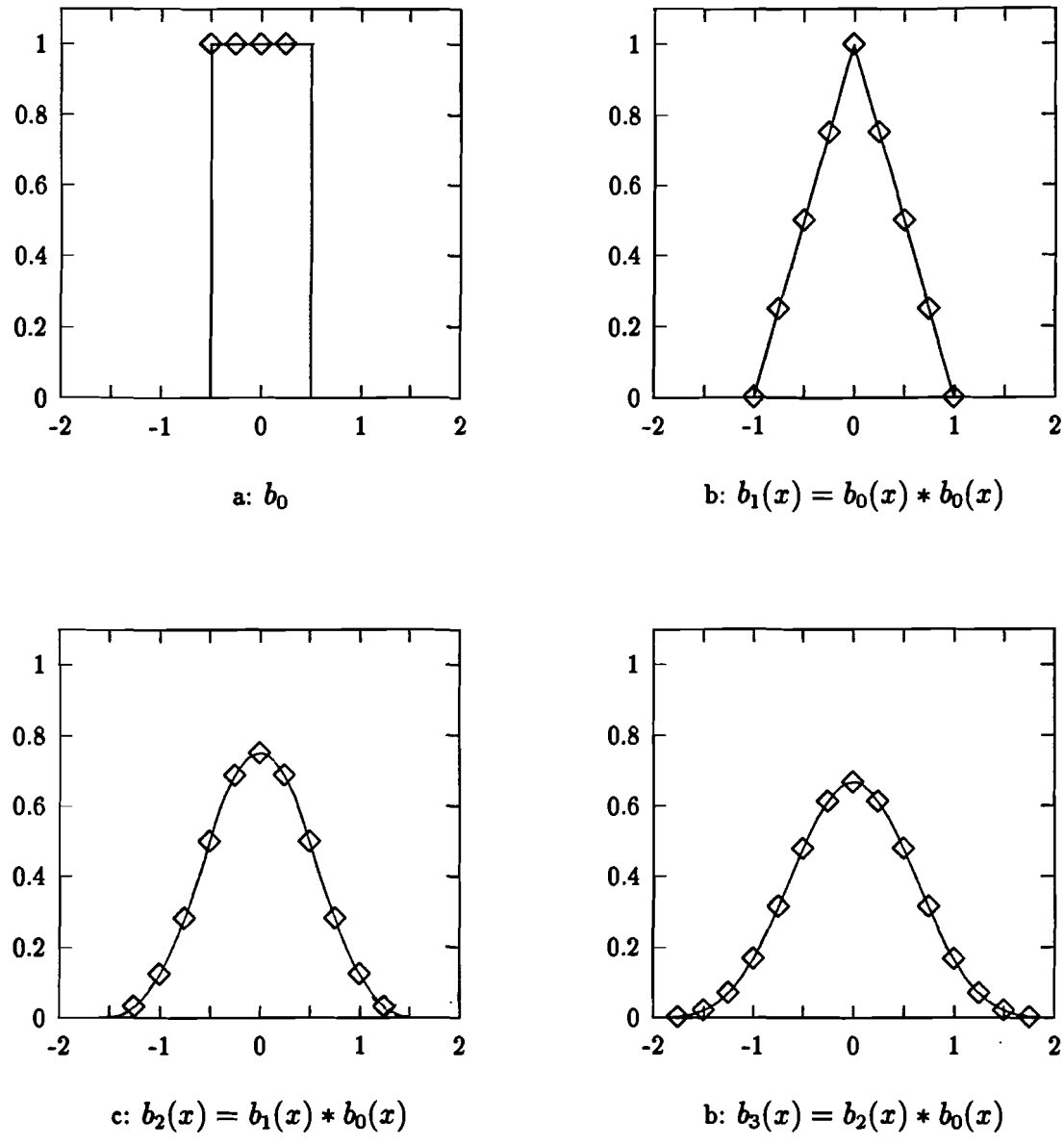


Figure 2.3: Some examples of B-splines. a: zero order B-spline, also known as the sample-hold function or nearest-neighbour interpolation function. b: first order B-spline, also known as linear interpolation function. c: second order B-spline and d: third order or cubic B-spline. The drawn curves are the continuous splines. The points in the curves are the factor four sampled discrete splines.

In figure 2.3 some B-splines are drawn. For the zero and first order B-splines the similarity with a gaussian function is rather abstract, but for the higher order splines it is undeniable. From the convolution property of B-splines (2.9) and the convolution property of gaussians (2.2) follows that the spread of the B-spline of order n is proportional to $\sqrt{n}$. For $n \geq 5$ a good approximation for the spread is found to be $\sigma = 0.433\sqrt{n}$. Furthermore, the B-splines have finite extent, their length is equal to their order plus one. So the objection to the gaussian function mentioned earlier is no longer valid.

The input image is densely sampled by a scanner with a sampling function that is approximately a delta function. Now we would like the interpolated image to have the same format, because then we can use the theory presented in the next chapter to determine the sampling filter. Therefore we represent the interpolation functions by discrete B-splines. These are simply derived from continuous B-splines through sampling them. The result is a discrete B-spline function of order n and with sampling factor m, given by

$$b_m^n(k) = \beta^n\left(\frac{k}{m}\right), \tag{2.10}$$

with k integer. When using a discrete B-spline to simulate a gaussian function with spread σ, the order of this discrete B-spline will be $n = (1/0.433)^2\sigma^2 = 5.33\sigma^2$.

Chapter 3

Sampling and interpolation

An image is a two-dimensional signal of finite extent. It can be represented by a vector in a Hilbert space, where the pixels in the image are the coordinates of the vector. Doing so we have a vector space model for image processing. In this model, the operations on images are represented by linear operators in the Hilbert space.

If some assumptions are made, the sampling and interpolation operators can be implemented using digital filters. The first assumption is that interpolation is performed the same way for each pixel. In that case, sampling can also be performed the same way for each pixel. The second assumption is that the signals are of infinite extent. Because in practice the signals are of finite extent, filtering gives distortions at the borders of the image. This distortion is eliminated by cutting off the edges of the image.

In this chapter it is described how the vector space model is used to describe sampling and interpolation of discrete signals with infinite sizes. The result is a generalized sampling theory. This theory is applied specifally to B-spline interpolation functions. We only discuss one-dimensional sampling and interpolation functions. Because the sampling points are positioned rectangularly, the two-dimensional filters are separable and filtering can be performed by successively filtering in the x- and y-direction.

3.1 The vector space model

Suppose we have a discrete signal $f(n), n \in \mathcal{Z}$. Now we want to sample this signal in a regular way, which means that we extract from it a signal $f_s(k)$, that is completely specified by the values $k = \ldots, -m, 0, m, 2m, \ldots$. We now introduce the set $\mathcal{D} = \{\ldots, -m, 0, m, 2m, \ldots\}$ and call the number m the sampling factor. After sampling we want to convert $f_s(k)$ back to a signal in $\mathcal{Z}$. This step is called interpolation, it is done by convolving the sampled signal with the interpolation

function $g(n)$. This results in the interpolated signal

$$\tilde{f}(n) = \sum_{k \in \mathcal{D}} f_s(k) \cdot g(n-k). \tag{3.1}$$

There are infinitely many ways to choose the sample values $f_s(k)$. The problem is how to choose them while guaranteeing that $\tilde{f}(n)$ is the best possible approximation of $f(n)$.

To solve this problem we describe the process of sampling and interpolation in terms of a vector space model. For a full description of this model is referred to [Oakley 1990, Martens 1992]. Only a brief summary of the results is given here.

In the model a signal $f(n)$ is represented by a vector $\boldsymbol{f}$ in a N-dimensional Hilbert space V with the orthonormal basis $B_\psi = \{\psi_n = \delta(n-o); n, o \in \mathcal{Z}\}$. In general a vector $\boldsymbol{f}$ can be written as a linear combination of its basis vectors, according to

$$\boldsymbol{f} = \sum_{n \in \mathcal{Z}} \hat{f}_\psi(n)\, \varphi_n. \tag{3.2}$$

The values $\hat{f}_\psi(n)$ are called the representation or coordinates of the vector $\boldsymbol{f}$ in the basis B_ψ. Thus for the vector space V the coordinates of a vector $\boldsymbol{f} = f(n)$ are equal to the function values of the signal. The vector product (inner product) of vectors $\boldsymbol{f}$ and $\boldsymbol{g}$ is defined as

$$\langle \boldsymbol{f}, \boldsymbol{g} \rangle = \sum_{n \in \mathcal{Z}} f^*(n)\, g(n). \tag{3.3}$$

The interpolated signal is an element of the M-dimensional subspace U of the Hilbert space V. The subspace has the repetitive basis $B_\varphi = \{\varphi_k; k \in \mathcal{D} = \{\ldots, -m, 0, m, 2m, \ldots\}\}$, with $\varphi_k = g(n-k)$, $n \in \mathcal{Z}$ and $k \in \mathcal{D}$. Now the basis vectors of U are all replicated versions of the interpolation function $g(n)$. The subspace has a dual basis B_ϕ, which is bi-orthogonal to B_φ. The Gramian matrix of the basis B_φ maps the coordinates of a vector in B_φ into the representation in B_ϕ. It is defined as the $N \times N$ matrix

$$\boldsymbol{G}_{\varphi\varphi} = [g_{\varphi\varphi}(i,j) = \langle \varphi_i, \varphi_j \rangle]_{i,j \in \mathcal{D}}. \tag{3.4}$$

Both vector spaces U and V are infinite dimensional. The signal $\boldsymbol{f}$ is an element of V. Through sampling we want to find the vector in U that is closest to $\boldsymbol{f}$. The vector in U which has minimum distance to a vector $\boldsymbol{f}$ in V is the perpendicular projection $P\boldsymbol{f}$ of that vector onto U. So sampling is performed by projecting the vector on a subspace. Interpolation is done by describing the resulting vector in U by its representation in the basis B_ψ. The complete process of sampling and interpolation is shown in figure 3.1.

The representation in B_φ of the perpendicular projection of $\boldsymbol{f}$ is

$$\hat{Pf}_\varphi(k) = \langle \phi_k, \boldsymbol{f} \rangle = \sum_{n \in \mathcal{Z}} \phi_k(n) f^*(n) = \sum_{n \in \mathcal{Z}} \phi(n-k) f^*(n). \tag{3.5}$$

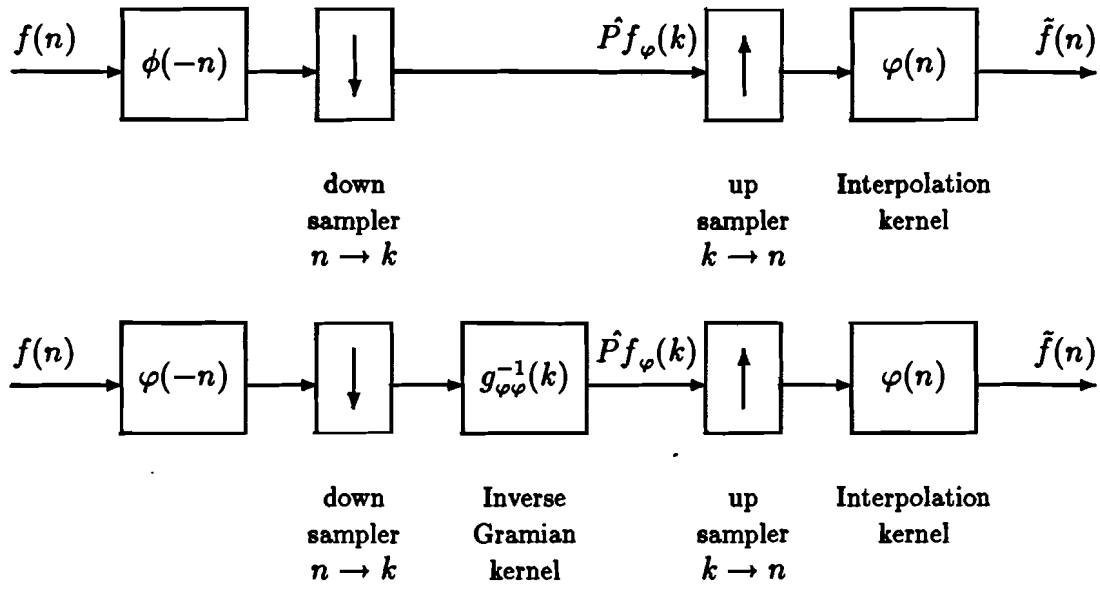


Figure 3.1: The complete sampling and interpolation process.

This projection is found by the first part of the process depicted in the upper diagram of figure 3.1. The function $\phi(n)$ is called the sampling kernel. Thus sampling is performed through filtering the input signal $f(n)$ with the mirrored sampling kernel, followed by down-sampling the result, i.e. measuring it at the sample points $k \in \mathcal{D}$. The projection operator in the lower diagram gives the representation in the dual basis B_ϕ,

$$\hat{P}f_\phi(k) = \langle \varphi_k, f \rangle = \sum_{n \in \mathcal{Z}} \varphi_k(n) f^*(n) = \sum_{n \in \mathcal{Z}} g(n-k) f(n). \tag{3.6}$$

This operation consists of convolving the input signal $f(n)$ with the mirrored interpolation function $g(-n)$, followed by down-sampling the result. The convolution is performed through filtering the signal with a filter with impulse response $g(-n)$. The inverse Gramian operator $G_{\phi\phi} = G_{\varphi\varphi}^{-1}$ transforms this representation into a representation in B_φ. This results in

$$f_s(k) = \hat{P}f_\varphi(k) = \sum_{l \in \mathcal{D}} g_{\phi\phi}(k-l) \hat{P}f_\phi(l), \tag{3.7}$$

which is equal to digitally filtering the downsampled signal with the inverse Gramian kernel $g_{\phi\phi}(k) = g_{\varphi\varphi}^{-1}(k)$. In the next section a method to invert the Gramian kernel is discussed. The Gramian kernel is found according to equation (3.4). Finally, interpolation is done by the last part of figure 3.1. This operator maps the coordinates $\hat{P}f_\varphi(k)$ of the projected signal to coordinates in the basis B_ψ of the original vector space V. This results in the equation (3.1).

Now let us take a look at the Shannon sampling theorem [Shannon 1949]. According to this theorem optimal sampling and interpolation is performed if we both sample and interpolate a signal with the kernel

$$\varphi(n) = \frac{\sin(\pi n/m)}{\pi n/m}. \tag{3.8}$$

Then the basis B_φ is orthonormal and the dual basis B_ϕ is equal to B_φ. This way the vector space model gives a generalization of the Shannon sampling theorem.

3.2 B-spline sampling and interpolation

Using a B-spline as the interpolation function, equation (3.1) changes into

$$\tilde{f}(k) = \sum_{l=-\infty}^{\infty} f_s(l) \cdot b_m^n(k - ml), k \in \mathcal{Z} \tag{3.9}$$

Some well-known examples of B-spline interpolation are nearest neighbour interpolation, which is performed with the zero order B-spline, and linear interpolation, performed with the first order B-spline. Also the cubic (third-order) B-spline, which has nice smoothing properties, is often used for interpolation [Hou 1978].

The basis, formed by the replicated zero order B-spline, is orthonormal. Thus, the Gramian kernel of this basis is a unity operator and the sampling filter is equal to the interpolation filter. For orders $n > 1$ it is obvious that this will not be the case, because the B-splines have a value greater then zero on neighbouring sample points. So down-sampling the signal without pre-filtering will not give optimal sample values.

The sampling filter's impulse response is equal to the convolution of the mirrored interpolation kernel and the inverse Gramian kernel. To calculate the inverse Gramian kernel first the Gramian kernel

$$g_{\varphi\varphi}(k) = \langle b_m^n(l), b_m^n(l-k) \rangle, \tag{3.10}$$

has to be determined. It is equal to the autocorrelation of the interpolation kernel. Because the convolution property (2.9) does not hold for discrete B-splines [Unser 1992], they have to be found by the straightforward calculation

$$g_{\varphi\varphi}(k) = \sum_{l=-\infty}^{\infty} b_m^n(l) b_m^n(l-k). \tag{3.11}$$

After that the inverse Gramian kernel

$$g_{\phi\phi}(k) = g_{\varphi\varphi}^{-1}(k) \tag{3.12}$$

has to be found. To do this we follow the method of [Unser 1991, 1993]. The Gramian kernel $g_{\varphi\varphi}(k)$ is characterized by its z-transform

$$\mathcal{Z}\{g_{\varphi\varphi}(k)\} = \sum_{i=-\infty}^{\infty} g_{\varphi\varphi}(i)\, z^i. \tag{3.13}$$

The z-transform of the inverse kernel is

$$\mathcal{Z}\{g_{\phi\phi}(k)\} = \mathcal{Z}\{g_{\varphi\varphi}(k)\}^{-1}, \tag{3.14}$$

which is equal to the polynomial fraction

$$\mathcal{Z}\{g_{\phi\phi}(k)\} = \frac{c_0}{(z^n + z^{-n}) + \left(\sum_{i=1}^{n-1} a_i \cdot (z^i + z^{-i})\right) + a_0}, \tag{3.15}$$

with

$$c_0 = \frac{1}{g_{\varphi\varphi}(n)} \quad \text{and} \quad a_i = a_{-i} = \frac{g_{\varphi\varphi}(i)}{g_{\varphi\varphi}(n)}.$$

We now find that

$$\begin{aligned} \mathcal{Z}\{g_{\phi\phi}(k)\} &= c_0 \prod_{i=1}^{n} \left(\frac{-z_i}{(1 - z_i z^{-1})(1 - z_i z)} \right) \\ &= c_0 \prod_{i=1}^{n} \left(\frac{z_i}{1 - z_i^2} \left(\frac{1}{1 - z_i z^{-1}} + \frac{1}{1 - z_i z} + 1 \right) \right), \end{aligned} \tag{3.16}$$

in which the z_i's are the roots of the polynomial formed by the product of the denominator of (3.15) and z^n. Because this polynomial is symmetrical, the roots z_i and z_{2n+1-i} are identical. Schoenberg has also shown that these roots are all simple and negative [Schoenberg 1969, lemma 8]. So only n roots have to be found. This is done with Laguerre's method starting the iteration in $z = 0$ [Press 1986, Kincaid 1991].

Finally the inverse kernel is found by taking the inverse z-transform of (3.16). This results in

$$g_{\phi\phi}(k) = c_0 \coprod_{i=1}^{n} \frac{-z_i}{1 - z_i^2}\, z_i^{|k|}, \tag{3.17}$$

in which $\coprod$ has the meaning of taking the n successive convolutions of the argument. These convolutions are found by straightforwardly calculating them. Because the power series $z_i^{|k|}$ have infinite length, they have to be truncated at a certain point. The length of the intermediate convolutions is chosen so that the error caused by the truncation is beneath a predetermined level.

The inverse Gramian kernel of a cubic B-spline with sampling factor $m = 4$ is shown as an example in figure 3.2. The sampling kernel, which is found by interpolating the inverse kernel with the factor four down-sampled cubic B-spline,

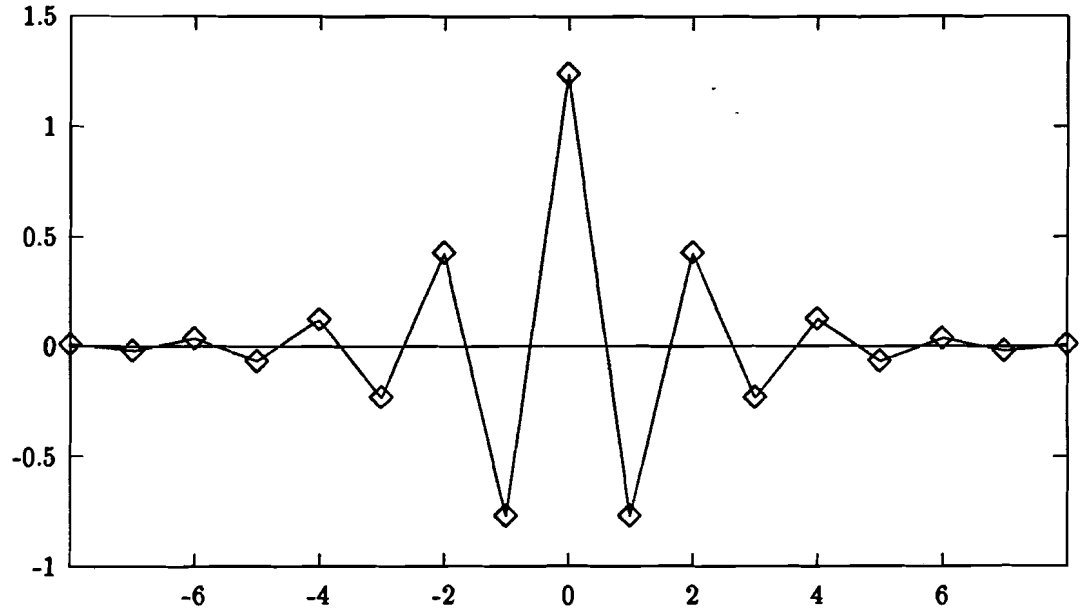


Figure 3.2: The inverse Gramian kernel for the discrete cubic B-Spline interpolation kernel with sampling factor $m = 4$.

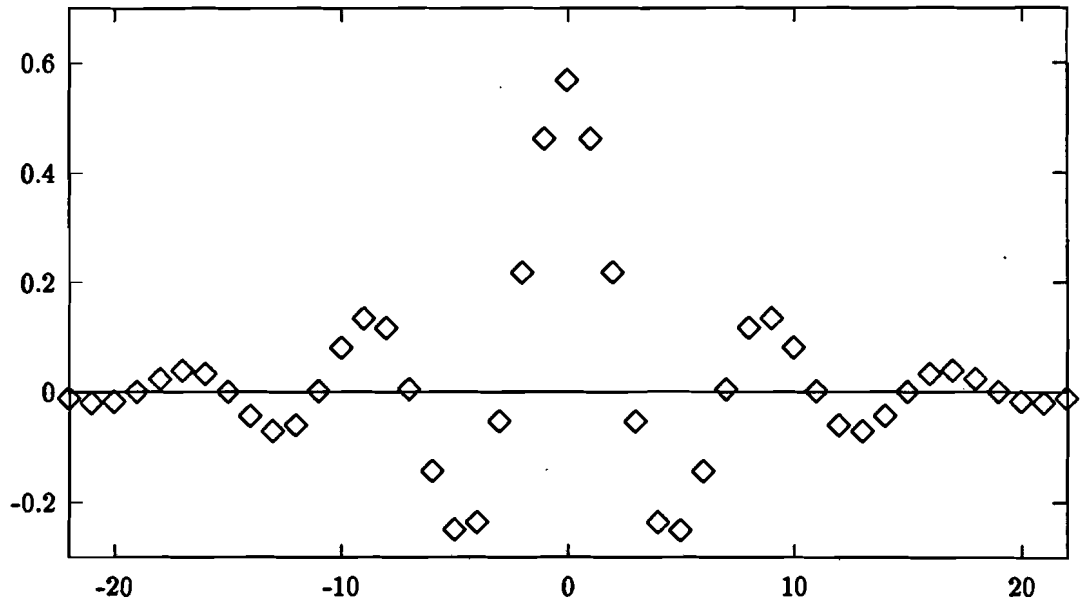


Figure 3.3: The sampling kernel for the discrete cubic B-Spline interpolation kernel with sampling factor $m = 4$.

is shown in figure 3.3. The inverse Gramian kernels all typically have alternating signs for each point k. As a result the sampling kernels show oscillatory behaviour and the higher the order n the slower the decay. This oscillatory behaviour is the cause of the artefact known as ringing. Suppose the signal is a mirrored step-function, the moment it falls, also the samples $[P\boldsymbol{f}]_\phi$ become zero. However, filtering with the inverse kernel causes the sample values $[P\boldsymbol{f}]_\varphi$ to oscillate for a while. In an image this oscillation will be visible along edges.

3.3 Frequency domain characteristics

Now we will take a brief look at the characteristics of the B-spline sampling and interpolation filters in the frequency domain. From sampling theory it is known that the sampling filter should cut off frequencies greater than half the sampling frequency (the Nyquist frequency) to prevent aliasing. In sampled images aliasing can cause disturbing visible patterns known as Moiré patterns, periodic structures which are not present in the original image [Legault 1973]. On the other hand perceptual tests have shown that it is better not to cut off all frequency components above the Nyquist frequency [Arguello 1982]. This is because the aliasing that is occasionally introduced this way is far less disturbing than the perceived loss of sharpness. Sharpness is strongly dependent of high-frequency components in the image.

In figure 3.4 the frequency spectra of the first and third order B-spline sampling and interpolation filters are shown. These spectra are obtained by simply calculating the FFT (Fast Fourier Transform) of the kernels shown in figure 2.3 and of their sampling kernels. Also the frequency spectra of the cascade of both filters, which are the products of their individual spectra, are shown.

Reitmeier has intuitively derived and tested some criteria for designing sampling and interpolation filters [Reitmeier 1981]. He has come to the following conclusions:

- The sampling filter should cut off frequencies above the Nyquist frequency to prevent aliasing. Best results are obtained if the actual cut-off frequency is slightly above the Nyquist frequency.

- The filter cascade should not have a too sharp cut-off. A sharp cut-off causes ringing, the introduction of shades parallel to strong contrast changes (edges) in the image.

From the power spectra, shown in figure 3.4, it can be seen that the B-spline sampling and interpolation filters behave this way. Two properties are especially noticeable. Firstly, the higher the B-spline order, the sharper the cut-off of the filter cascade. From this it can be expected that enlarging the order will increase the ringing artefact. Secondly, the sampling filter has a peak near the Nyquist

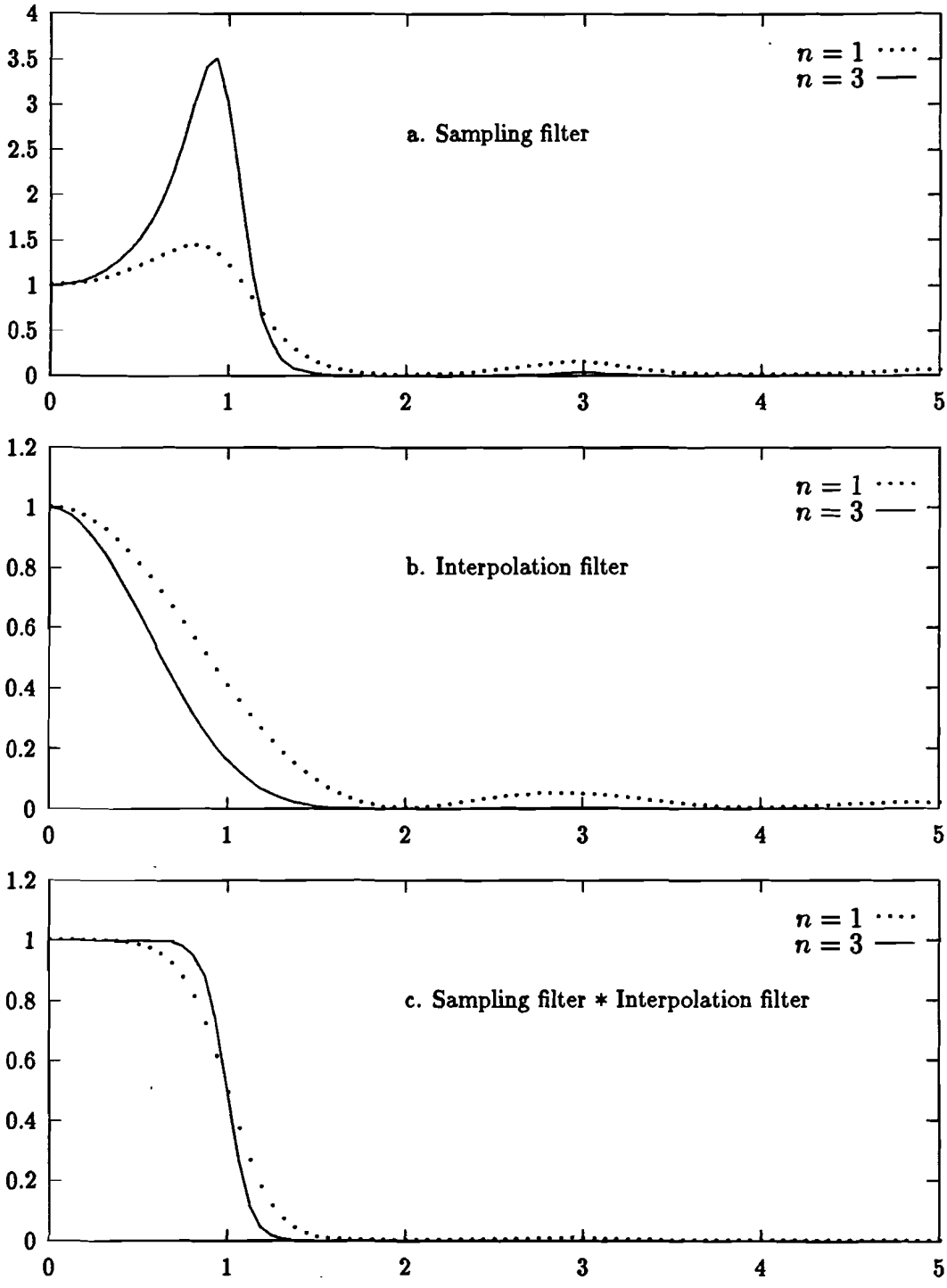


Figure 3.4: Power spectra of the first and third order B-spline sampling resp. interpolation kernels. In the last figure the product of both spectra is drawn. The frequency, on the horizontal axis, is expressed in terms of the samping frequency.

frequency. This peak compensates for the smooth decay of the interpolation filter in that region. So we expect the sampling filter to have the effect of sharpening the image. Ratzel already demonstrated that digitally filtering an image, which is sampled with a gaussian function, with a kernel that has high-pass characteristics improves the sharpness [Ratzel 1980, Schreiber 1985]. The vector space model presented in this chapter gives a mathematical justification for using such a kernel.

Chapter 4

Multidimensional scaling of sampling artefacts

Sampling and interpolating an image impairs it in multiple ways. From chapter 3 we expect the reproduced image to be blurred and to be degraded by ringing. Blur is a single and ringing is a complex perceptual attribute. In general, such perceptual attributes all have their influence on the perceived quality of an image. A way to model this influence is to describe quality as a linear combination of the sensational strengths of the perceptual attributes [Nijenhuis 1993].

Another approach is to describe a psychological space in which all stimuli are positioned. The way the stimuli are configured in this space determines how one perceives them. The distance between two stimuli is related to the perceived dissimilarity between the stimuli. Each perceptual attribute is related to a direction in the space. The projection of a stimulus on an attribute direction is linearly related to the sensational strength of that attribute in the stimulus. In this psychological space also directions can be found that are related to the quality and preference. The angle between these directions and the attributes' directions determine the influence of that attribute on the quality or preference.

If we are able to measure the perceived dissimilarities between all stimuli, we can construct the space. The directions can be found through scaling the sensational strengths of all attributes and finding a direction in the space that fits these strengths. This is all done in a multidimensional scaling (MDS) experiment. In this chapter an experiment by which the quality and the attributes blurring and ringing are scaled, is described. A review of the different methods of multidimensional scaling and their use can be found in [Kruskal 1977]. A brief summary of the models used in our multidimensional scaling experiment is included in appendix A.

Different assumptions can be made about the psychological space. We can assume it to be different for all persons or to be the same for everyone. This last model is denoted as homogeneity of perception. It means that all subjects

perceive the stimuli the same way. Individual preferences are not caused by different perceptions, but by different personal likes and dislikes of the attributes. In this case we try to find one single psychological space, common to all persons, in which the attribute directions can be different for everyone. This is proven to be a useful way for analyzing the perceived quality of images [Escalante]. The number of parameters is reduced to a minimum and the results can be easily interpreted.

If we want to do such a MDS experiment for sampled images, we have a problem. We want to measure the quality of the processed image. This means that we have to find a way to show both the original image, which is on the slide, and the processed image, which is on the display. Unfortunately it is not possible to show both under equal viewing conditions. The solution to this problem is to simulate the viewing situation. This is done by simulating overall interpolation functions with different spreads with B-spline interpolation functions, as will be demonstrated below.

In this chapter first the different stimuli and the experimental procedure are discussed. After that the results of the experiments are presented and the psychological space is reconstructed. Fitting the directions, special attention is paid to the sensitivity of the goodness of fit for a rotation of the direction. Finally some conclusions are drawn.

4.1 Experimental method

Because it is not possible to show both the original and sampled image under equal viewing conditions, the viewing situation is simulated. In this simulation, the scenes are not solely sampled, but also interpolated to their original sizes. Using a high resolution display[1], a part of the densely sampled image can be shown next to the corresponding of the sampled and interpolated image. Taking the densely sampled image as the reference, the quality of the sampled and interpolated image can be measured.

The different stimuli are produced through sampling and interpolating the scenes with B-spline functions of different orders. Taking the display function and the visual system into account, the interpolation functions can be interpreted as simulating the overall interpolation functions for different viewing distances.

The densely sampled images are only filtered by the display function and the visual system. These will have negligible spreads compared to the B-spline interpolation filters. So the densely sampled images can be seen as good approximations of the images on the slides.

Eight subject participated in this experiment. They all were related to the IPO and had experience in doing psychophysical scaling experiments. All subjects

[1]Barco CCID 7351B High-resolution 70 Hz interlace monitor, with a Gould DeAnza IP8400 image-processing system in the high-resolution mode.

had normal or corrected-to-normal vision.

4.1.1 The stimuli

The scenes used are WANDA and STADHUIS. These particular scenes are chosen because they have completely different characteristics. WANDA is a portrait of a woman. It contains merely smooth curves. STADHUIS is a picture of a old town hall. It contains high contrasts, sharp edges and a clearly visible brick structure. It is expected that this makes it easy to see the impairments caused by ringing and blurring. In figure 4.1 both scenes are shown.

The scenes have sizes of 1024×1024 pixels. They are obtained through down-sampling the 2048×2048 pixel output images of the scanner, with a sampling factor two. Because the sampling function of the scanner is by approximation a delta function, this results in the same images the scanner will produce if its output size is set to 1024×1024 pixels. Because the displayed scene STADHUIS was impaired by line-flicker, which would give an unwanted clue, it is decided to filter the scene with a 2×2 averaging filter.

The scenes are sampled and interpolated using different sampling and interpolation filters. The interpolation filters used are B-splines. For each interpolation filter the sampling filter is calculated according to the theory of chapter 3. Two parameters are changed to produce the stimuli; the sampling factor m and the order n. The used sampling factors are 2, 3 and 4 and the orders are 1, 5 and 9. In a pilot experiment the factor two sampled stimuli appeared to be too hardly distinguishable from the original scene, therefore only the one with $m = 2$ and $n = 1$ was included.

The result is a total of seven stimuli per scene. From these stimuli the edges are cut off, resulting in images with sizes of 500×944 pixels. Pairs of two stimuli are shown next to each other on the Barco monitor. The pixels and the peak-luminances of the dels are related according to the power law

$$L(x, y) \propto (f_s(x, y))^{\gamma_d}, \tag{4.1}$$

with $\gamma = 1.5$. This is done using a look-up-table. The monitor is placed in a dark room with a dimly lit white background. The viewing distance is taken to be three times the image height, which makes the visibility of the pixel structure just below threshold. To accurately control this small distance (0.75m for an image height of 0.25m) a head rest is used.

The B-spline interpolation functions simulate overall interpolation functions with different spreads. The spreads can be determined according to chapter 2. In the spread the non-linearity of the display will be taken into account. Because in the simulated situation the visual system and the display function filter both the processed and the reference image, they are not taken in account in the overall interpolation function. The spread of the overall interpolation filter now is equal

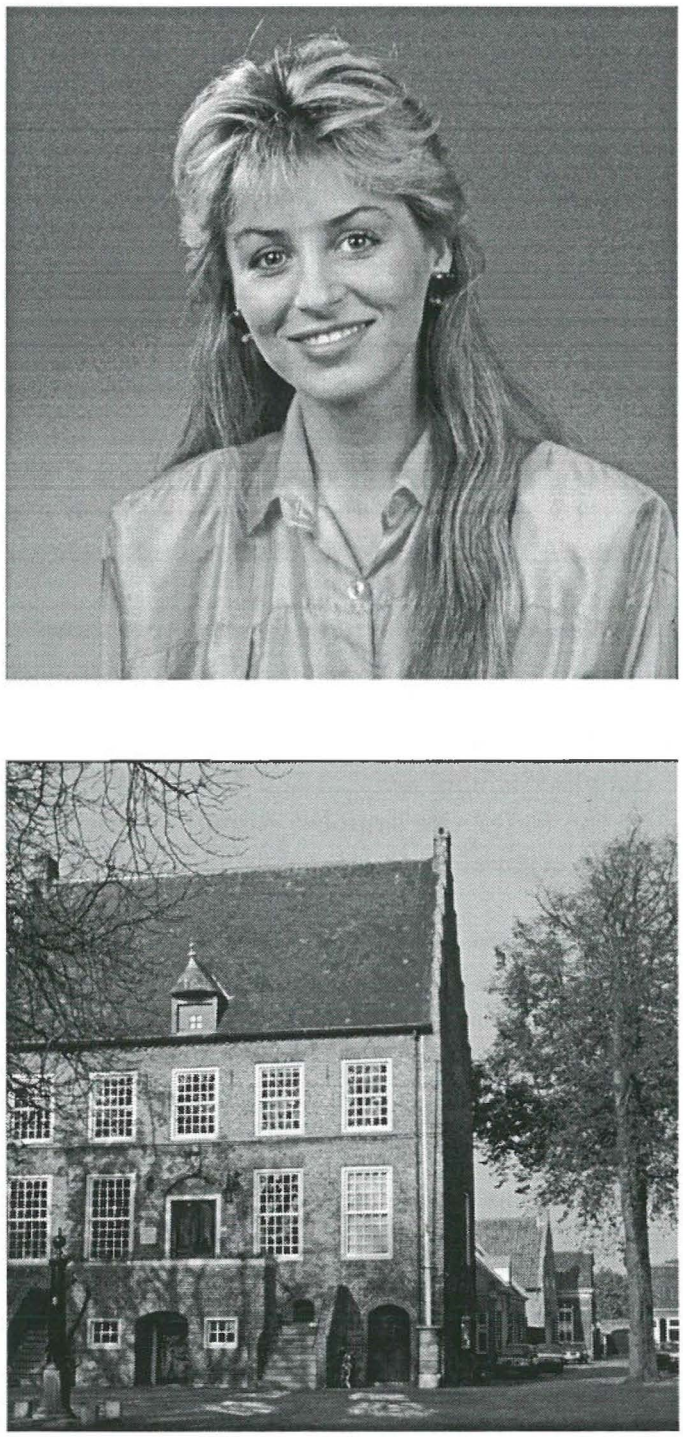


Figure 4.1: The scenes used in the experiments. The upper scene is WANDA, the lower is STADHUIS.

Table 4.1: The interpolation function's spread in min. of arc.

	$n=1$[2]	$n=5$	$n=9$
$m=2$	2.15	–	–
$m=3$	3.02	5.13	6.90
$m=4$	4.03	6.85	9.20

Table 4.2: The pixel distance in min. of arc.

	$n=1$	$n=5$	$n=9$
$m=2$	2.24	–	–
$m=3$	3.36	3.36	3.36
$m=4$	4.48	4.48	4.48

to the spread of the B-spline interpolation function. Expressed in terms of δ, the distance between two neighbouring dels, the spread is

$$\sigma_\delta = 0.433m\sqrt{n\gamma}. \tag{4.2}$$

If we express it as a function of the viewing angle θ, we find

$$\sigma_\theta = \arctan\left(\frac{\sigma_\delta}{3 \cdot 1024}\right). \tag{4.3}$$

In table 4.1 the values of σ_θ for the combinations of m and n are shown. The pixel distance of the sampled image is m times as large as the pixel distance of the sampled and interpolated image. Expressed in terms of the viewing angle, the pixel distance of the sampled image is

$$\delta_\theta = \arctan\left(\frac{m}{3 \cdot 1024}\right). \tag{4.4}$$

In table 4.2 this distance is shown.

The spreads cover a wide range. The reason for this is that the filters were chosen to make the impairments by the different attributes more distinguishable. If we would produce more realistic spreads, the sampling factor should be $m = 2$ and order n should lie between 0 and 5. In that case subjects would not be able to distinguish the stimuli from the original scene. To illustrate this the *peak signal-to-noise ratio* (PSNR) is calculated for all stimuli [Netravali 1988]. This ratio is defined as

$$\text{PSNR} = 10\log_{10}\left(\frac{A^2}{\text{MSE}}\right) \qquad \text{(dB)}, \tag{4.5}$$

with A as the peak value of the stimulus, which is $A = 255$. In this relation the mean squared error

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(f(i) - \tilde{f}(i)\right)^2, \tag{4.6}$$

[2]For this B-spline order equation (4.2) does not give a good approximation of its spread. A better approximation is $\sigma = 0.57m\sqrt{\gamma}$. This one is actually used in the table.

Table 4.3: The PSNR of the stimuli

	A	B	C	D	E	F	G
	$m=2$ $n=1$	$m=3$ $n=1$	$m=3$ $n=5$	$m=3$ $n=9$	$m=4$ $n=1$	$m=4$ $n=5$	$m=4$ $n=9$
WANDA	34.8	31.9	32.5	32.4	30.3	30.2	30.5
STADHUIS	32.4	28.5	29.3	29.3	26.6	26.8	26.7

in which N is the number of pixels, is used. Here $f(i)$ are the pixels of the original scene and $\tilde{f}(i)$ are pixels of the sampled and interpolated image. The PSNR's of the stimuli are shown in table 4.3. We see that the stimuli produced with the B-splines with $n = 1$ and $m = 2$ have a high PSNR. The perceptual threshold is about 35 to 40dB, images with a higher PSNR are not distinguishable from the original scene. For example the PSNR of a 8 bit quantized image is

$$\text{PSNR} = 10.8 + 6R = 10.8 + 6 \cdot 8 = 55.8 \qquad \text{(dB)}, \tag{4.7}$$

which is below threshold [Netravali 1988]. So we have to use at least a sampling factor of 3 to make the impairments clearly perceivable.

4.1.2 Procedure

The experiment consisted of four sessions. In the first session the subjects were asked to scale the dissimilarity between all combinations of two stimuli numerically. The seven stimuli of each scene were completed by a 500 × 944 pixel part of the densely sampled scene. The subjects were presented a pair of two images, next to each other. Each pair was shown twice, the second time both images were interchanged. This resulted in a total of 64 image pairs per scene. They were presented in random order, alternately a pair of the scene WANDA and the scene STADHUIS. Each pair was shown during approximately 20 seconds, after that the subject was asked to give his judgment on a ten-points scale from zero to nine. To adjust the sensitivities of the subjects' scales, this sequence was preceded by a trial sequence of eight image pairs. In this sequence each of the sixteen stimuli was included. The pairs were chosen so that the dissimilarities which were expected to be the extremes were included in the trial sequence.

In the second session again the question was to scale the dissimilarity between two stimuli, which were shown next to each other. Only now each time one of the stimuli was the densely sampled scene, which served as a reference. So in this experiment the quality of the reproduction is scaled. The fourteen processed stimuli were shown in random order, six times each, three times at the left of the original scene and three times at the right. The experiment sequence was preceded

by a trial sequence of ten stimuli. Again each pair of images was shown during approximately 20 seconds and the subjects were asked to give the judgment on a scale from zero to nine.

In the final two sessions the fourteen stimuli were presented one at a time, in random order. During the experiment again each stimulus was repeated six times and the experiment was preceded by a trial sequence of ten images. Each image was shown during approximately 12 seconds. In the third session the subjects were asked to scale the sharpness of the image and in the fourth the amount of ringing. To clarify the concept ringing, a demonstration image was shown before the fourth session. This was the image STADHUIS with a size of 1024×1024 pixels. It was sampled and interpolated with a ninth order B-spline with subsampling factor four.

The impairments by ringing and blurring are demonstrated in the pictures in figure 4.2. The picture above is a heavily blurred WANDA. Blurring strongly affects the visibility of detail in an image. The image below shows a detail of the demonstration image mentioned above. The impairment by ringing is visible as shades along the wall.

4.2 Experimental results

The results of the two parts of the experiment are discussed separately. First the results of the scaling experiments are presented. Then the psychological space with the attribute directions is constructed and discussed.

4.2.1 The scaling experiments

Results

The subjects all did three scaling experiments, one in which the quality of the reproduction is scaled, one in which the sharpness is scaled and one in which the impairment by ringing is scaled. The categorical scaling results of the subjects are converted into sensational strengths on a psychological scale according to Thurstone's law of categorical judgment [Thurstone 1927]. The individual results obtained this way are depicted in appendix B. From this appendix can be seen that the results are fairly homogeneous over all subjects. In the conversion condition D is used, which means that the interval borders are assumed to have no spread and all sensational strengths have the same spread [Torgerson 1958]. Using this last assumption all individual sensational strengths are divided by their respective spreads. This transformed the scale so that the spread of the averaged sensational strength of each stimulus becomes equal to one. After that the results are averaged over all subjects. This gives the averaged results depicted in figure 4.3. The characters on the horizontal axis indicate the sampling and interpolation filters used, the same way as in table 4.3. The scale on the vertical

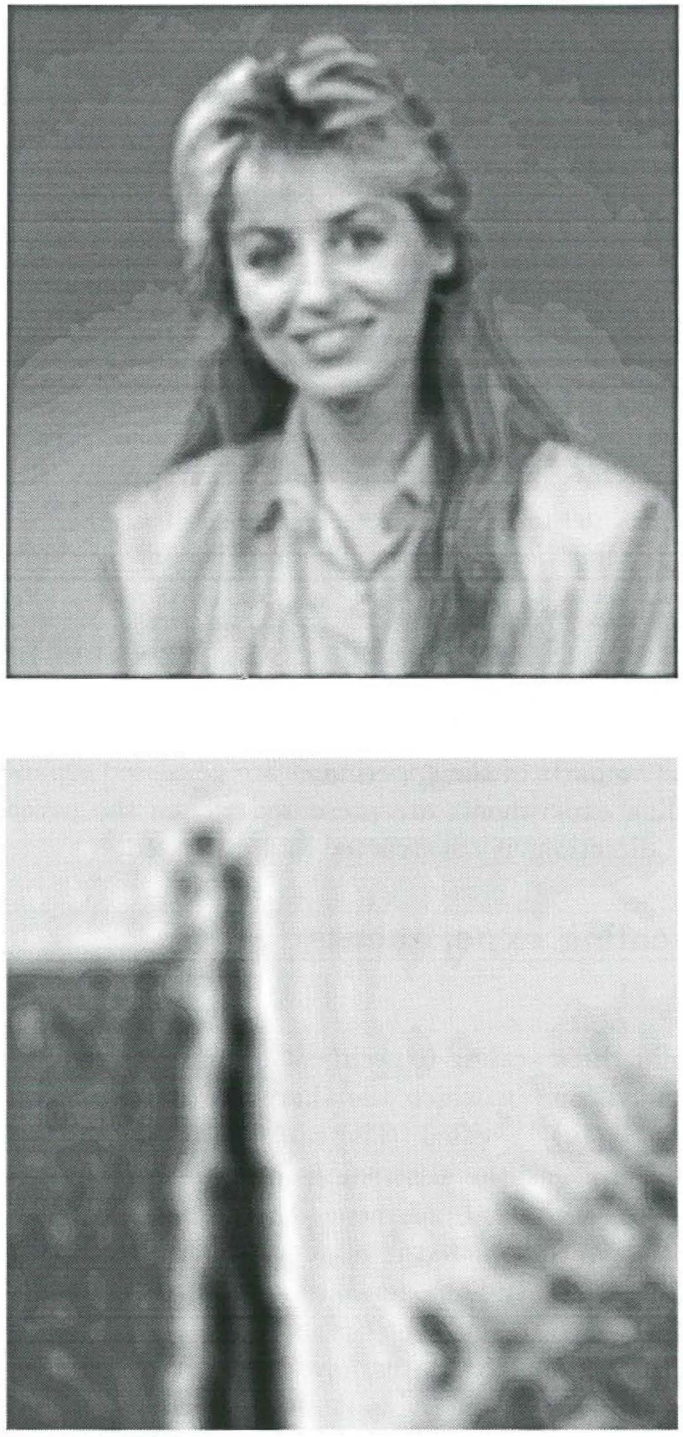


Figure 4.2: Demonstration of the impairments by blurring and ringing. The picture above is blurred and in the picture below ringing is visible.

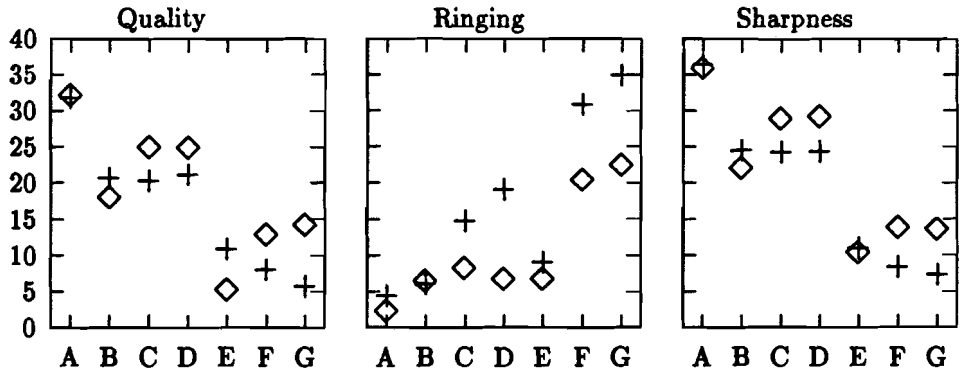


Figure 4.3: The averaged results of the attribute scaling experiments. The symbol ◇ indicates the scene WANDA and the symbol + the scene STADHUIS. The characters on the horizontal axis indicate the sampling and interpolation filters used, the same way as in table 4.3.

axis is obtained by dividing the individual results by their respective spreads. Because there were eight subjects, the spread in the resulting figures is equal to $\sigma = 1/\sqrt{8}$.

Discussion

The stimuli with different sampling factors can easily be distinguished. For an increasing sampling factor m the preference decreases, impairment by ringing gets stronger and sharpness decreases too. However, this is not surprising. When we throw more information away, we can expect the reproduction to become worse.

The linearly interpolated stimuli ($n = 1$, $m = 3$ and $m = 4$) are impaired by a staircase effect, which is a visibility of sampling structure. In appendix B can be seen that some subjects judge the amount of ringing in these stimuli very high. It appears that they mix up the staircase effect and ringing. Furthermore, for the scene WANDA these stimuli have a lower quality judgment, because there the staircase effect is very disturbing.

The quality judgments for the scene WANDA are higher than the judgments for the scene STADHUIS. This is in agreement with the PSNR's in table 4.3, from these figures also follows that WANDA is less degraded by sampling and interpolation than STADHUIS.

Finally sampling and interpolation with higher order B-splines does not affect the quality significantly.

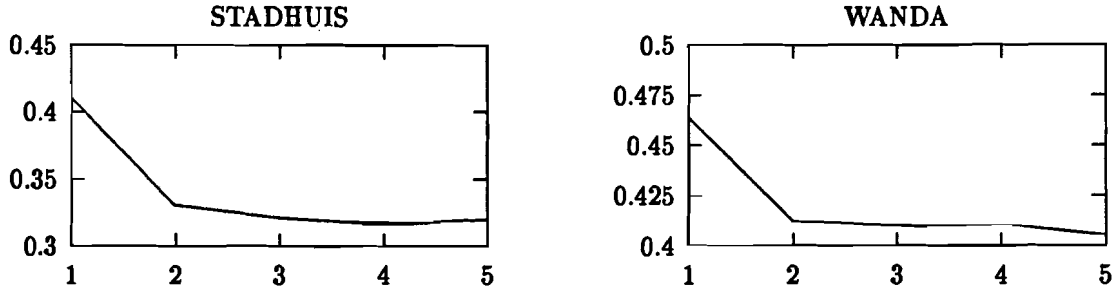


Figure 4.4: The maximum likelihood standard error estimate (vertical axis) for different dimensionalities of the psychological space (horizontal axis).

4.2.2 The dissimilarity experiment

Results

The dissimilarity scaling results are used to construct the psychological space for each scene. This is done with the multidimensional scaling program MULTISCALE. With this program it is possible to choose between several different models for the way the dissimilarity data is scaled, the way the error distribution is modeled, etc. The used models are explained in appendix A. The input of the program is a half-matrix without diagonal, with the dissimilarities. We obtain this matrix by averaging the two dissimilarity judgments that each subject gave to each stimulus pair.

The constructed psychological space is assumed to be two-dimensional. There are two arguments to motivate this assumption. The first is that monitoring the standard error estimate of the configuration for different dimensions, a two-dimensional configuration is found to give good results. In figure 4.4 the value of the error is plotted. If a larger dimension is taken, the configuration does not get significantly better. The second argument is that a two-dimensional configuration is easy to interpret. It is impossible to draw higher dimensional configurations in one plot.

Both configuration are drawn in figures 4.6 and 4.7. Tables 4.5 and 4.6 show the within-subject correlations. The horizontal axis in the configuration denotes the direction that explains most of the variance in the dissimilarity data. Although the within-subject correlations are not all equally good, there is no reason to eliminate subjects from the analysis. Especially the configuration of the scene STADHUIS fits the dissimilarity judgments of the subjects well. The ellipses in the figures denote the 95 percent confidence regions. If all stimuli except one are kept in position, there is a 95 percent probability that the variable stimulus lies within the ellipse surrounding its position.

The individual attribute vectors are found using the multiple linear regression

Table 4.4: The directions for the attributes quality (Q), ringing (R) and sharpness (S) for each subject. The last column shows the common attribute direction. The figures denote the angle between the direction and horizontal axis.

scene WANDA

	1	2	3	4	5	6	7	8	Com
Q	-5.2	35.7	-17.6	-17.3	-20.7	13.7	10.1	-8.1	-7.2
R	79.8	94.2	101.5	100.9	96.1	114.0	95.5	103.8	100.0
S	-19.2	-42.9	-42.4	-16.7	-31.1	-18.2	-16.5	-15.4	-21.6

scene STADHUIS

	1	2	3	4	5	6	7	8	Com
Q	-35.6	-12.0	-39.3	-34.8	-26.8	-17.2	-36.6	-9.5	-25.5
R	146.1	173.1	120.1	142.1	143.6	189.2	151.6	149.2	-29.8
S	-38.7	-43.5	-30.9	-25.0	-26.4	-7.4	-21.5	-31.9	151.7

model described in appendix A. For this model the sensational strengths are taken as its input. The resulting directions are expressed by their angles with the horizontal axis in table 4.4. The ringing attribute direction originally gave a very bad fit, because visibility of sampling structure was mixed up with ringing by some subjects. Therefore the two stimuli with $n = 1$ and $m = 3$, $m = 4$ are left out of the analysis. This improved the fit significantly. Tables 4.5 and 4.6 show the within-subject correlations between the sensational strengths and the projections on the associated direction in the psychological space.

For these directions the sensitivity of the goodness of fit is determined. This is done through calculating how the attribute judgments are correlated with the projections on different directions. These calculations indicate that the direction can be varied over a wide range without affecting the goodness of fit very much. In figure 4.5 this is illustrated for two attributes. The horizontal axis in this figure denotes the direction, expressed in the angle between the direction and the horizontal axis of the psychological space. The vertical axis shows the correlation between the attribute judgments and the projection on the directions, i.e. the goodness of fit. We see that the plots typically have a flat top. Especially if the best-fitting direction is about horizontal, the top is very wide. This flat tops show that the goodness of fit is quite insensitive for a change in the direction.

These sensitivities suggest that for each attribute one single direction can be found, that perhaps fits all subjects. These common directions for the attributes quality (Q), ringing (R) and sharpness (S) are found using the model in

Table 4.5: Within-subject correlations for scene WANDA

Configuration

Subject	1	2	3	4	5	6	7	8
	0.76	0.73	0.81	0.84	0.91	0.86	0.91	0.71

Individual attribute directions

Subject	1	2	3	4	5	6	7	8
Quality	0.988	0.959	0.930	0.931	0.941	0.957	0.977	0.964
Ringing	0.941	0.903	0.931	0.888	0.939	0.827	0.981	0.930
Sharpness	0.896	0.976	0.981	0.996	0.970	0.993	0.993	0.938

Common attribute directions

Subject	1	2	3	4	5	6	7	8
Quality	0.987	0.871	0.926	0.927	0.933	0.939	0.964	0.964
Ringing	0.941	0.895	0.930	0.888	0.936	0.792	0.975	0.926
Sharpness	0.896	0.945	0.952	0.995	0.965	0.993	0.992	0.936

Table 4.6: Within-subject correlations for scene STADHUIS

Configuration

Subject	1	2	3	4	5	6	7	8
	0.87	0.93	0.82	0.89	0.83	0.95	0.82	0.86

Individual attribute directions

Subject	1	2	3	4	5	6	7	8
Quality	0.926	0.956	0.923	0.929	0.968	0.941	0.903	0.969
Ringing	0.958	0.987	0.957	0.970	0.976	0.972	0.965	0.968
Sharpness	0.957	0.969	0.962	0.956	0.966	0.975	0.960	0.930

Common attribute directions

Subject	1	2	3	4	5	6	7	8
Quality	0.921	0.951	0.912	0.925	0.968	0.939	0.896	0.960
Ringing	0.957	0.965	0.817	0.965	0.972	0.918	0.964	0.968
Sharpness	0.951	0.953	0.962	0.956	0.966	0.959	0.958	0.929

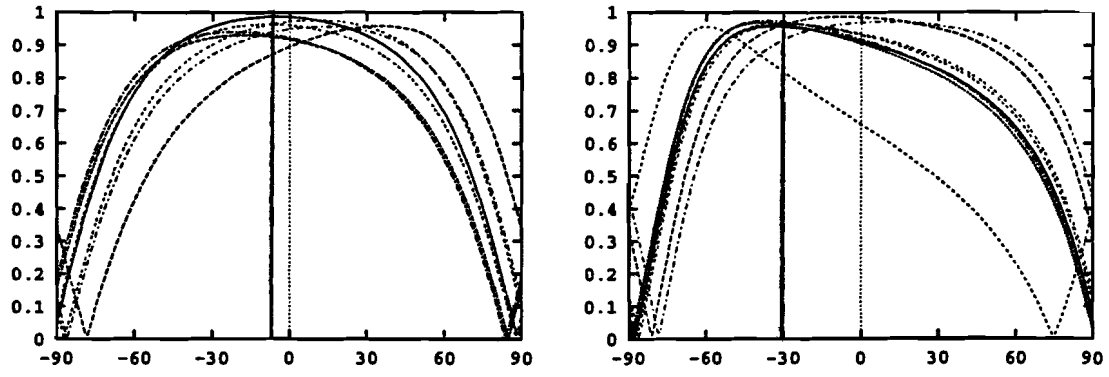


Figure 4.5: The sensitivity of the correlation for a change in the direction. The horizontal axes denote the directions, for the figure on the left this is the direction related to the quality of the scene WANDA and for the figure on the right this is the direction of the ringing in the scene STADHUIS. The vertical axes denotes the correlation. For each subject a curve is drawn.

appendix A. They are also shown in table 4.4 and in figures 4.6 and 4.7. From the within-subject correlations of these common directions, shown in tables 4.5 and 4.6, can be concluded that they fit the scaling data almost as good as the individual directions.

Discussion

Stimuli with the same sampling factor cluster in the psychological space. This is in agreement with what we saw in the results of the quality scaling experiment discussed above. The confidence regions show that there is some uncertainty about the positions of the stimuli within the clusters, but that the clusters are clearly separated.

For the scene WANDA the ringing direction is almost orthogonal to the quality direction. The sharpness direction is almost equal to the quality direction. This means that for this scene the quality judgment is not affected by the ringing attribute, but is almost completely determined by the sharpness of the image.

4.3 Conclusions

Sampling and interpolating an image with higher order B-splines does not affect the quality significantly, as is demonstrated by the clusters of stimuli with the same sampling factors. When the order is enlarged, the sharpness of the image stays the same while the strength of the attribute ringing increases. In scenes like

STADHUIS, that contains sharp edges, ringing is better visible than in smooth scenes like WANDA.

Sharpness is the most important attribute of an image. Ringing is not that disturbing, one subject even reported that he had not noticed it, till he was asked to scale it. This suggests to test a sampling filter, that increases both the strength of the attributes ringing and sharpness. Because the strength of the ringing in an image does not have a strong influence on the quality judgment, we can expect the quality to get better in that case. In a next experiment we will test this expectation.

If we look at the direction of the attribute ringing for the scene STADHUIS, we see that when the stimuli denoted by F and G are projected on this direction, their order is interchanged, compared to the original scaling results. The correlation still is very high, but this is due to the small amount of stimuli that is used determining the direction. If the stimuli denoted by B and E, which were not included in the optimization of the direction, are projected on the ringing direction, the strength of the ringing in these stimuli is too large. These two observations lead to the conclusion that the direction of the attribute ringing can not be fit in the constructed psychological space. Possibly the psychological space should be of a higher dimensionality. However, with the used stimuli, it is not possible to find an additional dimension from the dissimilarity scaling results. That the attributes lie on one line in this configuration, does not mean that the psychological space is one-dimensional. There are some attributes that are not scaled, like the staircase attribute.

Ringing is often mixed up with the staircase effect. The staircase effect is an interpolation attribute. It appeared because linear interpolation is used. In an actual viewing situation, where the sampled images are not interpolated with low-order B-spline functions, this attribute will not be present. Therefore we have not paid much attention to it.

The pixel distances of the sampled images are larger than in usual viewing situations. The experiment can also be done with a variable viewing distance to make the pixel distance equal to 1 min. of arc, for each sampling factor.

Instead of finding individual attribute directions for each subject, we may find common attribute directions in the psychological space, that fit all subjects. Doing so we can confine the multidimensional perception model. Then not only homogeneity of perception is assumed, but also is assumed that all persons have the same understanding of the attributes. This assumption is not always met [Escalante], but in our case it works. It is the experimenter's job to ensure that the subjects understand the meaning of the experiment and that they all are scaling the same attributes.

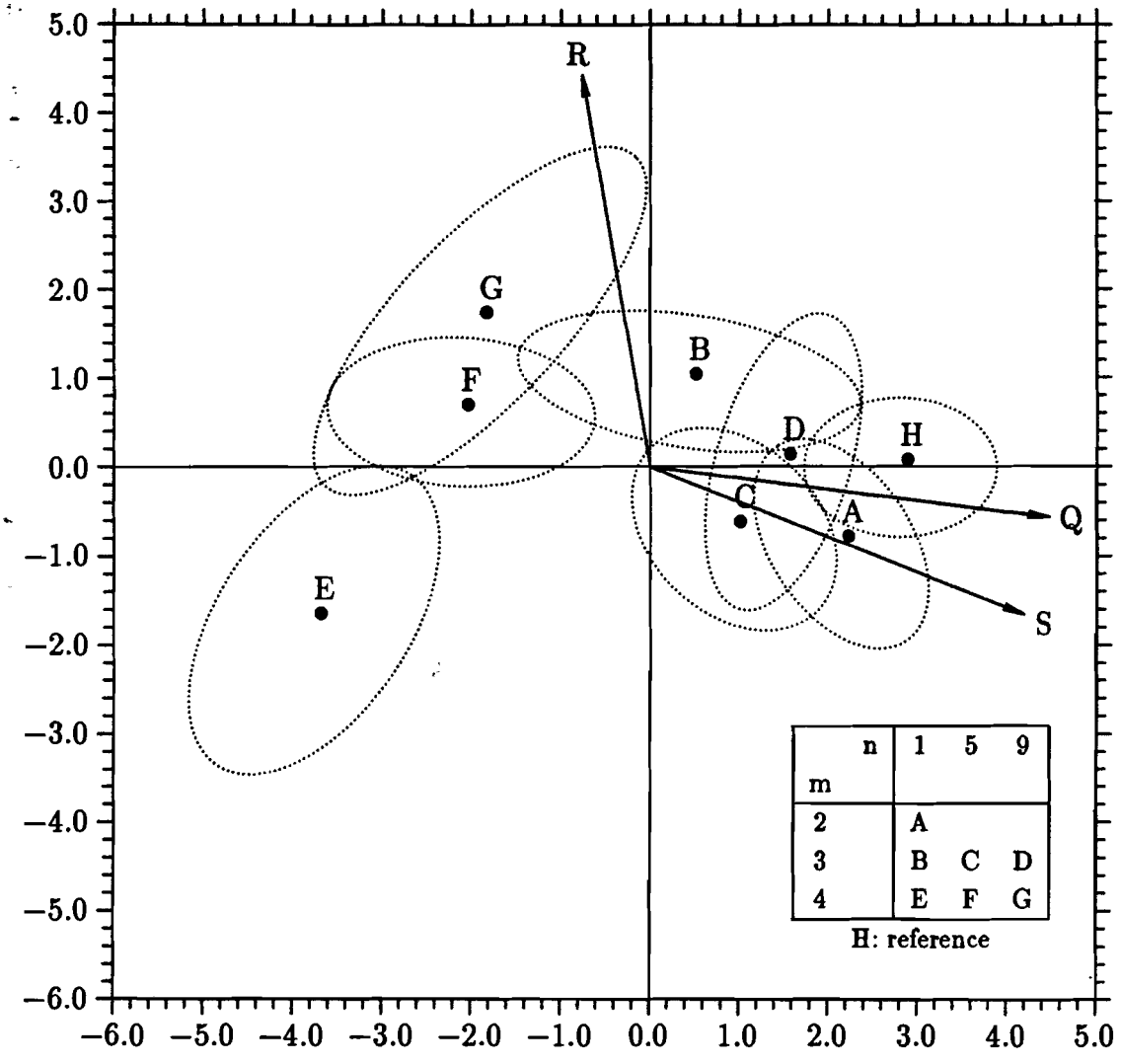


Figure 4.6: The MULTISCALE configuration of the scene WANDA, with the common attribute vectors and the 95 percent confidence regions. The horizontal axis denotes the first dimension, this dimension explained most of the variance in the data. The vertical axis denotes the second dimension.

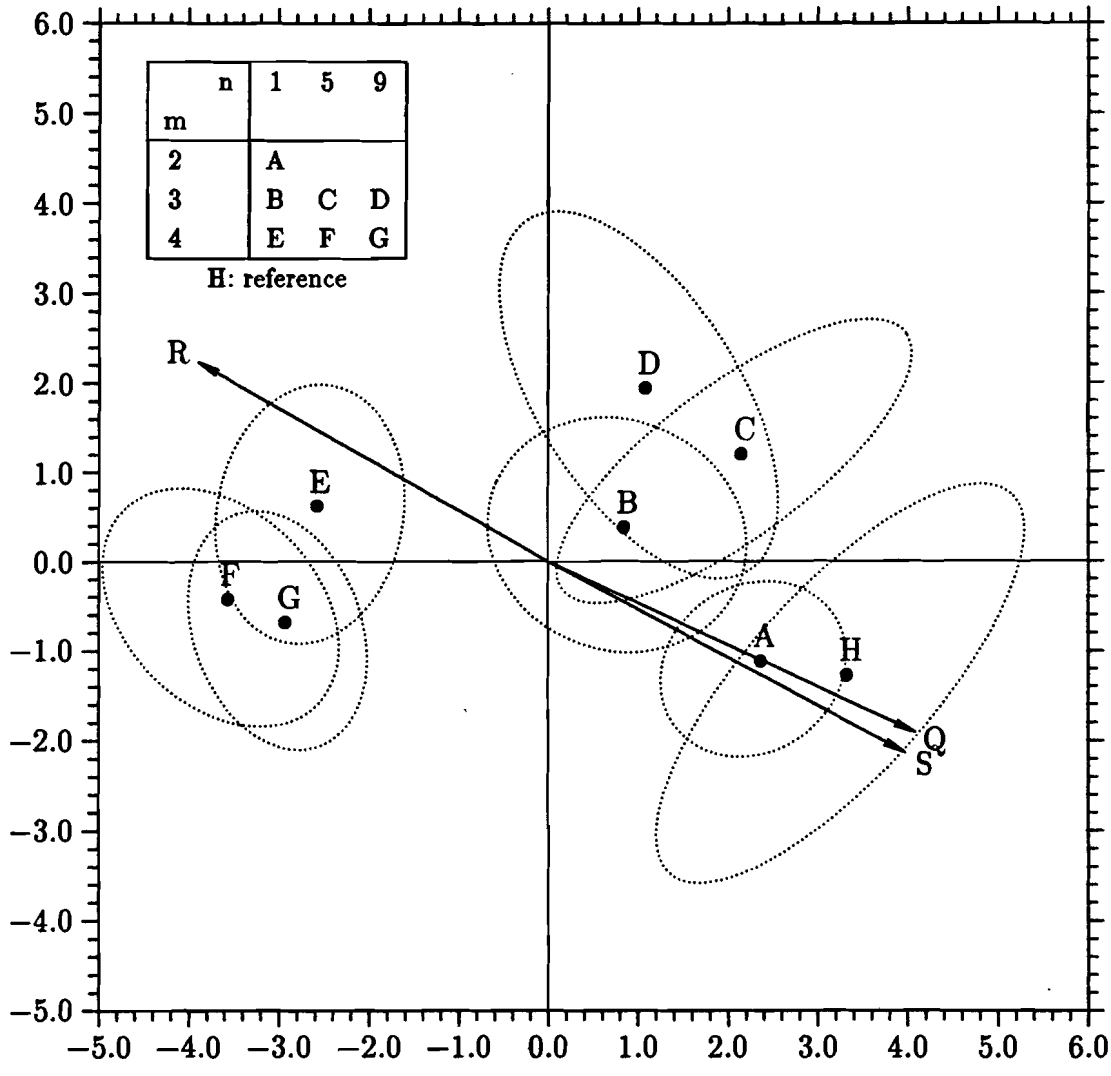


Figure 4.7: The MULTISCALE configuration of the scene STADHUIS, with the common attribute vectors and the 95 percent confidence regions. The horizontal axis denotes the first dimension, this dimension explained most of the variance in the data. The vertical axis denotes the second dimension.

Chapter 5

Testing different sampling functions in a simulated viewing situation

Now we will test different sampling filters for different interpolation filters. Like in the previous chapter, the viewing situation is simulated. This is done by simulating overall interpolation functions with different spreads by B-spline functions.

Again, for each interpolation filter the optimal sampling filter is determined. In addition, the sampling filters, designed for interpolation filters with larger spread than the actual interpolation filter's spread, are tested. From the observations on the frequency spectra of B-spline sampling filters, in chapter 3, it can be expected that these filters make the reproduction sharper but also increase the ringing attribute. From the conclusions of the previous chapter we now expect that this increases the image quality. Additionally, the reproduction produced without applying a sampling filter is tested. This is done to measure the profit of filtering prior to sampling. And finally, gaussian sampling filters are often used and said to give satisfactory results [Schreiber 1985]. Therefore, images also are sampled with a sampling kernel equal to the interpolation kernel.

A subject can be asked to scale the quality of an image, which means that he can compare every image with the original, and he can be asked to scale the preference of an image, which means that he has no reference. We expect both tasks to give different results. For when a sampling filter makes the reproduction look sharper than the original image, the preference may be high, while the quality actually is not good.

If the simulation of the viewing situation from the previous chapter is used, a comparison between both scaling tasks can be made. So the experiments described in this chapter have a twofold purpose. Firstly, different sampling filters are compared, and secondly, the relation between quality scaling and preference scaling is examined.

5.1 Experimental method

The experiment consisted of two sessions. In the first session the preference is measured, in the second the quality. In both experiments six subjects participated. They also participated in the MDS experiment. The experimental situation was equal to the experimental situation in the previous chapter. Again the scenes are sampled and interpolated, simulating different viewing situations. But because this time not all combination of two images have to be shown, more stimuli can be used in the experiment and different sampling functions can be compared.

5.1.1 The stimuli

For the reasons mentioned in the previous chapter, the scenes WANDA and STADHUIS are used. They have sizes of 1024 × 1024 pixels. To eliminate line-flicker in the displayed STADHUIS scene, it is processed by a 2 × 2 averaging filter. The scenes are sampled and interpolated using different sampling and interpolation functions. Four different interpolation functions are used. These are B-splines in which the sampling factor m is either of 2 or 3 and the order n is 1 or 5. The interpolation functions with $m = 4$ and $n = 9$ are not used in this experiment, because the stimuli they produced were too heavily impaired by these interpolation functions in combination with the used sampling functions.

With each interpolation function, five sampling functions are used. The first is a delta function, this means no filtering at all. The stimuli produced this way are equal to the images the scanner would produce if its output size is set to 512×512 pixels. These stimuli are included to measure the profit of filtering prior to sampling. The second sampling function is made equal to the interpolation function. For small spread this method is often used. The third sampling filter is calculated according to the theory of chapter 3. These are the sampling functions used in the MDS experiment. If the interpolation function has order n then the fourth sampling function is the, according to chapter 3, optimal sampling kernel of the $(n+1)$-th order B-spline interpolation kernel. The fifth sampling function is the optimal sampling kernel of the $(n+2)$-th order B-spline interpolation kernel. These latter two filter are expected to improve the sharpness of the image, at the cost of adding more ringing.

This results in a total of twenty stimuli per scene. The peak signal-to-noise ratios of these stimuli is given in table 5.1. According to these PSNR's, the third sampling filter should give the best results. However, from the conclusions of the previous chapter, we expect the latter two stimuli to give better results.

From the stimuli the edges are cut off, resulting in images with sizes of 500 × 944 pixels. They are shown on a Barco CCID7351B monitor. A look-up-table is used to have the pixels and the peak-luminances of the dels related according to

Table 5.1: The PSNR of the stimuli

WANDA

m	2	2	3	3
n	1	5	1	5
1	31.5	30.2	30.9	28.7
2	33.3	29.4	30.5	27.3
3	34.8	35.7	31.9	32.5
4	31.8	34.5	29.1	30.2
5	28.1	31.1	27.1	28.8

STADHUIS

m	2	2	3	3
n	1	5	1	5
1	28.3	26.8	27.4	25.1
2	30.3	25.9	26.8	23.8
3	32.4	34.2	28.5	29.3
4	28.9	32.5	25.5	26.6
5	25.4	28.4	23.6	25.1

the power law

$$L(x,y) \propto (f_s(x,y))^{\gamma_d}, \tag{5.1}$$

with $\gamma = 1.5$. The monitor is placed in a dark room with a dimly lit white background. Because the viewing distance was very small (0.75m for an image height of 0.25m), a head rest is used.

5.1.2 Procedure

In the first session of the experiment the preference was scaled. The subjects was shown one image at a time. They were asked to give a quality judgment on a ten-point scale from zero to nine. In the second session the unprocessed image was also shown as a reference, each time at the left of the sampled and interpolated image. The subjects were asked to scale the quality of the reproduction, the same way they did in the MDS experiment. In both experiments the image was shown for approximately 10 seconds. All stimuli were repeated four times. Both experimental sequences were preceded by a trial sequence of sixteen stimuli, which were chosen to be a representative selection from the forty stimuli.

5.2 Experimental results

Results

The categorical scaling results of the subjects are converted into sensational strengths on a psychological scale, according to Thurstone's law of categorical judgment condition D [Thurstone 1927, Torgerson 1958]. In appendix C the individual results obtained this way are shown. In this appendix can be seen that the results are fairly homogeneous over all subjects. Therefore it is allowed to divide these results by their spreads and average them, the same way as the attribute scaling results in the previous chapter. The averaged results are shown

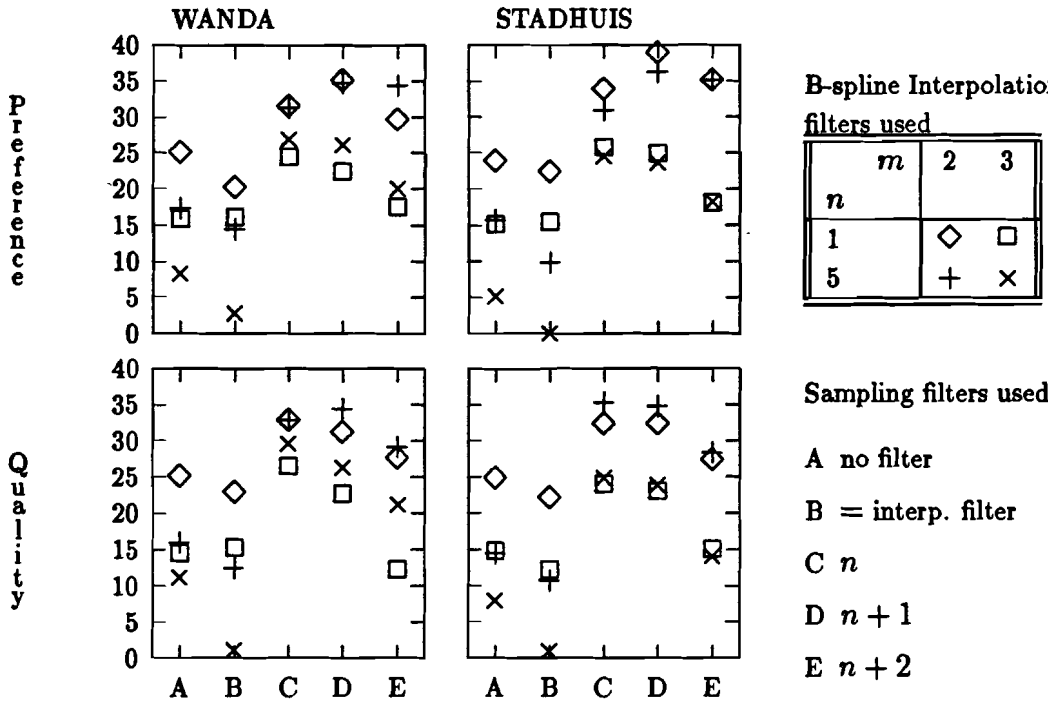


Figure 5.1: The averaged results of the quality and preference scaling experiments. The symbols ◇, □, + and × indicate the results for the different interpolation kernels according to the table next to the figures. The enumeration below this table indicates the meaning of the characters on the horizontal axes of the figures.

in figure 5.1. Next to the figures the meaning of the different symbols and the meaning of the characters on the horizontal axis are explained. The scales on the vertical axes originate from dividing the individual results by their respective spreads. Because there were six subjects, the spread of each stimulus in the resulting figures is equal to $1/\sqrt{6}$.

Discussion

In the figures we see that both quality and preference of the gaussian and delta sampling filters are worse than the quality and preference of the B-spline sampling filters. The results of the gaussian and delta sampling filters are acceptable for the first order interpolation filters, but not for the fifth order interpolation function. In that case the stimuli are heavily blurred[1].

For the B-spline sampling filters the order of the interpolation filter does not

[1]In fact, the blurred reproduction of WANDA in figure 4.2 is obtained by both sampling and interpolating WANDA with the same fifth order B-spline

have a significant influence on the quality and preference judgments. When the sampling factor increases, the quality decreases. These two observations are in accordance with the results of the previous chapter.

There is a clear difference between the preference and quality judgments. The $(n+1)$-th order sampling filters give the best preference. Also the $(n+2)$-th order sampling filters give good preference results. For the quality the n-th and $(n+1)$-the order sampling filters give approximately the same result and the $(n+2)$-th order filter performs much worse. Thus the higher order sampling filters do have the effect of making the reproduction look sharper than the reference image.

5.3 Conclusions

Making the sampling filter equal to the interpolation filter gives bad results for higher order interpolation filters. It appears to blur the image heavily. It should however be noticed that in literature this kind of sampling filters only is used in combination with interpolation filters with small spread ($\sigma_\delta/\delta < 1.0$). In that case ($n = 1$) we also obtain acceptable results.

We get a better reproduction if we optimize the sampling function for the interpolation function. If we overestimate the spread of the interpolation function in the optimization, we obtain the best results. This can be explained with two arguments. Firstly, doing so we make the reproductions sharper and increase the strength of the ringing. However, quality is mostly determined by sharpness, therefore this overestimation improves the quality. Secondly, the sampling filter is optimized for a discrete interpolation function, while the overall interpolation function actually is analogue. These two arguments demonstrate the well-know fact that the mean squared error perceptually is not a good error criterion.

There is a difference between the quality and the preference judgments. This is clearly visible for the $(n+2)$-th order B-spline sampling filters. This filter makes the reproduction sharper than the original scene. Though this reproduction looks better than the other ones, because it is the sharpest, it differs from the original scene and the quality is lower. The difference between both judgments is that for the quality the maximum is at a lower order.

Chapter 6

Testing different sampling functions in a real viewing situation

In a final experiment, sampling functions are tested in a more realistic viewing situation. This means that the scenes are sampled and then shown on the display. Interpolation is done by the display function and the visual system. If the assumption that the overall interpolation function can be modeled by a gaussian function is correct, then the B-spline sampling kernels should improve the image quality.

Again different sampling filters with different sampling factors and orders are tested. Only this time it is not possible to compare the image shown with a reference image. So only the preference can be measured. For drawing conclusions about the quality of the images, we have to rely on the comparison between quality and preference, made in the previous chapter.

6.1 Experimental method

In this experiment the same six subjects as in the previous experiment participated. The experiment was done between both sessions of the previous experiment. This way is prevented that subjects remembered the 'original' scenes of the quality scaling experiment in the previous chapter, which would bias the judgment.

6.1.1 The stimuli

The used scenes are again WANDA and STADHUIS, obtained by the slide-scanner. These are both sampled and displayed, resulting in stimuli with sizes

of 512 × 512 pixels. The used sampling factors are 2 and 4. For the factor two sampling functions, the sizes of the scenes are 1024 × 1024 pixels. For the factor four sampling functions the sizes of the scenes are 2048 × 2048 pixels. The 1024 × 1024 pixel scenes are obtained by down-sampling the 2048 × 2048 pixel scenes with a factor two. Because the scanner has approximately a delta function as its sampling function, it would produce the same scenes if its output size was set to 1024 × 1024 pixels.

Five different sampling functions are used for both scenes and both sampling factors. The first is a delta function, which is equivalent to scanning the analogue slide directly to a 512 × 512 discrete image, without filtering. Of course the resulting stimuli are equal for both sampling factors. They are not both included in the experiment. The other four functions are the sampling kernels of B-spline interpolation functions of order 0, 1, 2 and 3. Higher order B-spline sampling functions are not included, because they impair the scenes too much.

The result is a total of 18 stimuli. These are all shown on a Barco CCID7315B monitor, using a Gould DeAnza IP8400 image-processing system in the low-resolution mode to display them. A look-up-table is used to have the pixels and the peak-luminances of the dels related according to the power law

$$L(x,y) \propto (f_s(x,y))^{\gamma_d}, \tag{6.1}$$

with $\gamma = 2.5$. The monitor is placed in a dark room with a dimly lit white background. The viewing distance was equal to 5 times the image height, this was approximately 1.40m. This distance is a compromise between visibility of sampling structure and visibility of impairment by ringing and blurring. We do not want visibility of sampling structure, but we do want the subject to be able to distinguish between the stimuli. Unfortunately, as a result, the line structure was just visible.

6.1.2 Procedure

The six subjects were asked to give a quality judgment of the stimuli presented, on a ten point scale from zero to nine. Each stimulus was repeated eight times. The experiment was preceded by a trial sequence in which each stimulus was included once. The stimuli were shown eight seconds. Between two stimuli an adaptation field was shown for two seconds.

6.2 Experimental results

Results

The scaling results are processed the same way as in the preceding chapters. Using Thurstone's law of categorical judgment condition D, the judgments are

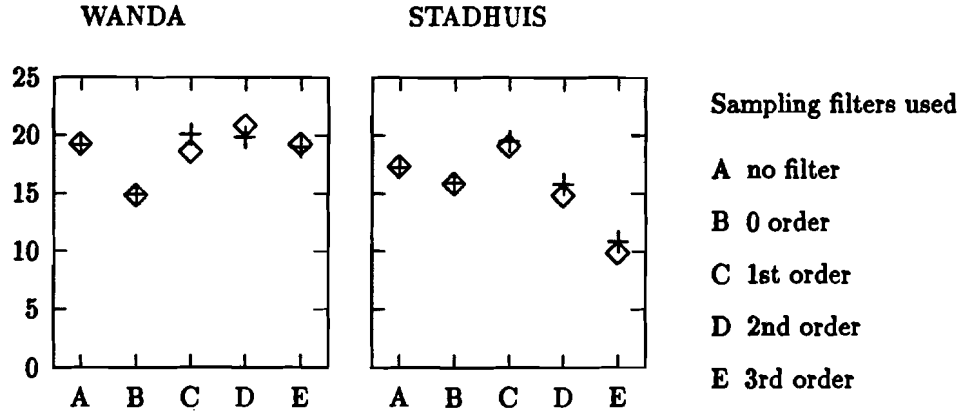


Figure 6.1: The averaged results of the experiment. The symbol ◇ indicates the factor two sampled stimuli and the symbol + the factor four sampled stimuli. The meaning of the characters on the horizontal axes is shown in the table next to the figure.

converted into sensational strengths [Thurstone 1927, Torgerson 1958]. In appendix D the individual results are shown. These results are fairly homogeneous over all subjects. Therefore it is allowed to divide these results by their spreads and average them. In figure 6.1 the averaged results are shown. Next to the figure the meaning of the characters on the horizontal axes is explained. The scales on the vertical axes originate from dividing the individual results by their respective spreads. Because there were six subjects, the spread in this figure is equal to $\sigma = 1/\sqrt{6}$.

The experiment was very difficult, for the differences between the images were hard to scale. Therefore it may be that the trial sequence was too short. Possibly, subjects did not have the opportunity to adjust the sensitivity of their scale well enough, with the result that the first judgments of each stimulus have a great deviation from the final result. Indeed, the results of some subjects and stimuli show this adjustment problem. Therefore, the scaling results are also processed with the first two judgments of each stimulus excluded. However, the adjustment problems were not that serious, because this treatment gave no improvement of the variance. In the results presented, all eight judgments are taken in account.

Discussion

The stimuli of WANDA all have a high quality, except for those that were filtered by an averaging filter (zero order spline). For STADHUIS the quality drops very strong for the second and third order filtered stimuli (D and E). This is also caused by line-flicker that was visible in these stimuli. We have not paid any attention to line-flicker yet, because like aliasing it is very scene dependent.

6.3 Conclusions

The sampling factor has no significant influence on the preference. Even when the scenes are down-sampled, without filtering prior to sampling, the resulting images still are of a high quality. This demonstrates that the scanner used has very low noise. The gain of scanning the images at a high resolution and sampling the result, is not very high.

From the previous chapter we know that when quality is compared to preference the maximum shifts towards a lower order. In the experiment described above we measured preference. For the scene WANDA there is a maximum if the sampling function is of order $n = 2$. For the scene STADHUIS the maximum is at $n = 1$, but we do not know whether the sharpness increases for higher order sampling filters because these stimuli were impaired by ringing. We now can conclude that for a maximum quality images should be sampled with a first order B-spline sampling filter.

The scenes sampled with an averaging filter have a low preference. In practice this sampling filter is often used [Netravali 1988]. The reason to use it is to eliminate aliasing and to decrease the amount of noise in the sampled image. The motivation given is that detail smaller than the sampling distance is not visible in the sampled image anyway. From chapter 3 we know that this sampling filter is only optimal when sample-hold interpolation is used.

With B-splines it is not possible to generate interpolation functions with random spread, the spread only can have discrete values ($\sigma = 0.433\sqrt{n}$). Therefore it is not possible to determine the exact location of the maximum using B-spline functions. However the stimuli are hard to distinguish anyway and a more exact location of the maximum is impossible to find with the approach we used.

Chapter 7

General conclusions

The vector space model tells us how to find the optimal sampling function for a given interpolation function, according to a minimum mean squared error criterion. Sampling an image with this optimal sampling function has the effect of preserving the sharpness in the image. Through psychophysical experiments it is proven that this model gives perceptually relevant results. However, the minimum mean squared error is not an adequate model of the perceived impairments in an image. If the sampling function is determined with this model, it is better to overestimate the spread of the interpolation function. This way the peak signal-to-noise ratio of the reproduction of the scene is lower, but the reproduction is sharper and therefore has a higher quality.

Sharpness is the most important attribute when sampling and interpolating images. The other examined attribute, ringing, does not have a large influence on the image quality. Especially in images without sharp edges, like WANDA, impairment by ringing is hardly visible. For the scene STADHUIS, the influence of ringing on the dissimilarity scaling results was so small, that it was not possible to find a direction in the constructed psychological space, that was accurately related to the strength of the impairment by ringing.

There is a clear difference between quality and preference judgments of an image. If there is no reference image shown, people have a higher preference for images that are sharper than the reference image.

In the experiments with the simulated viewing situation, the pixel distances of the sampled images are larger than in usual viewing situations. These experiments can also be done with a variable viewing distance to make the pixel distance equal to 1 min. of arc, for each sampling factor.

If images have to be shown on a display, they must be sampled. Best results are obtained if images are filtered with a first order B-spline sampling filter prior to sampling. This gives better results than filtering with a averaging filter (zero order B-spline) or gaussian filter, two approaches which are often used. However, the gain of sampling the image with such a sampling filter is not very high. If

the image is just scanned to a discrete image with the desired size, the quality is still very good. Hence, for the used scenes the influence of aliasing is limited and the scanner produces very low noise.

In the multidimensional scaling experiment described in this report, the goodness of fit of the attribute directions are in a wide range insensitive for changes in the directions. It is possible to find for each attribute one common direction that fits all subjects. Therefore we can safely assume that in our case all subjects have the same understanding of the attributes. This may generally be the case, though sometimes persons label the attributes differently. This assumption should be further tested in other multidimensional scaling experiments.

Bibliography

[Arguello 1982] R.J. Arguello, H.B. Kessler and H.R. Sellner. "The effect of sampling, optical transfer function shape, and anisotropy on subjective image quality," *Opt. Eng.*, Vol. 21, no. 1, pp. 23-29, 1982.

[Blommaert 1981] F.J.J. Blommaert and J.A.J. Roufs, "The foveal point spread function as a determinant for detail vision," *Vision Research*, vol. 21, pp. 1223-1233, 1981.

[Dubois 1985] E. Dubois, "The sampling and reconstruction of time-varying imagery with application in video systems," *Proc. IEEE*, Vol. 73, no. 4, pp. 502-522, 1985.

[Escalante] B. Escalante-Ramírez, J.B. Martens and H. de Ridder, "Multidimensional characterization of the perceptual quality of noise-reduced computed tomography images," To be published.

[Hou 1978] H.S. Hou and H.C. Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. ASSP-26, no. 6, pp. 508-517, 1978.

[Hummel 1983] R. Hummel, "Sampling for spline reconstruction," *SIAM J. Appl. Math.*, Vol. 43, no. 2, pp. 278-288, 1983.

[Kincaid 1991] D. Kincaid and W. Cheney, *Numerical analysis.* Pacific Grove, California: Brooks/Cole Publ. Comp., 1991, pp. 100-104.

[Kruskal 1977] J.B. Kruskal and M. Wish, *Multidimensional scaling.* The Sage series on methodology in social sciences. Bevely Hills, California: Sage publications, 1977.

[Legault 1973] R. Legault, "The aliasing problems in two-dimensional sampled imagery," in *Perception of displayed information*, L.M. Biberman, Ed. New York: Plenum Press, 1973, chapter 7.

[Martens 1992] J.B. Martens, *Image technology.* Instituut voor Perceptie Onderzoek, Eindhoven, syllabus, 1992.

[Netravali 1988] A.N. Netravali and B.G. Haskell, *Digital pictures, Representation and compression.* New York: Plenum Press, 1988.

[Nijenhuis 1993] M.R.M Nijenhuis, "Sampling and interpolation of static images: a perceptual view," Technische Universiteit Eindhoven, Ph.D dissertation, 1993.

[Oakley 1990] J.P. Oakley and M.J. Cunningham, "A function space model for digital image sampling and its application in image reconstruction," *Comput. Vision Graphics Image Process.*, Vol. 49, pp. 171-197, 1990.

[Peterson 1962] D.P. Peterson and D. Middleton, "Sampling and reconstruction of wave-number-limited functions in N-dimensional euclidian spaces," *Informat. Contr.*, Vol. 5, pp. 279-323, 1962.

[Press 1986] W.H. Press, et al., *Numerical recipes.* Cambridge: Cambridge University Press, 1986, pp 259-269.

[Ratzel 1980] J.N. Ratzel, "The discrete representation of spatially continuous images," Massachusetts Institute of Technology, Ph.D dissertation, 1980.

[Reitmeier 1981] G.A. Reitmeier, "The effects of analog filtering on the picture quality of component digital television systems," *SMPTE Journal*, Vol. 90, pp. 949-955, 1981.

[Schoenberg 1946] I.J. Schoenberg, "Contribution to the problem of approximation of equidistant data by analytic functions," *Quart. Appl. Math.*, Vol. 4, pp. 45-99, 112-141, 1946.

[Schoenberg 1969] I.J. Schoenberg, "Cardinal interpolation and spline functions," *J. Appprox. Theory*, Vol. 2, pp 167-206, 1969.

[Schreiber 1985] W.F. Schreiber and D.E. Troxel, "Transformation between continuous and discrete representations of images: a perceptual approach," *IEEE Trans. Patern Anal. Mach. Intell.*, Vol. PAMI-7, no. 2, pp. 178-186, 1985.

[Shannon 1949] C.E. Shannon, "Communications in the presence of noise," *Proc. IRE*, Vol. 37, pp. 10-21, 1949.

[Thurstone 1927] L.L. Thurstone, "A law of comparative judgment," *Psychological review*, Vol. 34, pp. 273-286, 1927.

[Torgerson 1958] W.S. Torgerson, *Theory and methods of scaling.* New York: Wiley, 1958.

[Unser 1991] M. Unser, A. Aldroubi and M. Eden, "Fast B-spline transforms for continuous image representation and interpolation," *IEEE Trans. Patern Anal. Mach. Intell.*, Vol. PAMI-13, no. 3, pp. 277-285, 1991.

[Unser 1992] M. Unser, A. Aldroubi and M. Eden, "Polynomial spline signal approximations: filter design and asymptotic equivalence with Shannon's sampling theorem," *IEEE Trans. Inform. Theory*, Vol. IT-38, no. 1, pp. 95-103, 1992.

[Unser 1993] M. Unser, A. Aldroubi and M. Eden, "B-spline signal processing," *IEEE Trans. Signal Processing*, Vol. 41, no. 2, pp. 821-848, 1993.

Appendix A

Multidimensional scaling

A multidimensional scaling experiment can be split into two parts. The first part is constructing the psychological space, i.e. finding the stimulus configuration. This is done doing a dissimilarity scaling experiment. If there are I stimuli then all I^2 possible combinations of two stimuli are presented to all subjects. They are asked to scale the dissimilarity between both stimuli. By means of these dissimilarity judgments the space can be constructed.

The second part is finding the directions of the perceptual attributes in the psychological space, this part is called the vector model. The model is found by doing for each attribute an experiment in which the subjects are asked to scale the strength of that attribute. The results of these experiments are used to optimize the attribute directions. In this appendix both parts are explained in more detail.

A.1 Constructing the stimulus configuration

Let I indicate the number of stimuli and R the number of subjects. The dissimilarity judgment of subject r for stimulus pair (i, j) will be denoted by d_{ijr} and the corresponding distance by $\tilde{d}_{ij}$. If the psychological space is N-dimensional, then x_{in} is the n-th coordinate of stimulus i.

The stimuli are positioned in such way that the Euclidean distance between two stimuli is related to the perceived dissimilarity between these stimuli. This relation does not have to be linear, i.e. the scaling is allowed to be ordinal. In fact, the logarithm of the distance between a stimulus i and a stimulus j, $\log \tilde{d}_{ij}$, will be fit to $f_r(\log d_{ij})$. Here $f_r(\log d_{ij})$ is

$$f_r(\log d_{ij}) = p_r \log d_{ij} + v_r, \tag{A.1}$$

which is a power transformation on the dissimilarity data.

The subjects' judgments are assumed to be representative for the complete population. This means that the dissimilarity judgments shall not differ from the

judgments resulting from an experiment done with a very large group of subjects, say the complete dutch population. The error distribution in the hypothetical population of possible judgments on a specific pair is assumed to be known. We assume the logarithm of this error to be normally distributed. In psychophysics this is referred to as Weber's Law. If we now have a random configuration, then we can calculate the likelihood that the distances in this configuration are produced by the dissimilarity data. This figure is called the likelihood estimation. By maximizing the logarithm of it, the configuration is optimized.

In practice the configuration will never fit the transformed dissimilarities perfectly, because of the noise on the data and the incompleteness of the model. The error is expressed by two figures. The within-subject standard error σ_r^2 is defined as the variance

$$\sigma_r^2 = \text{var}(d_{ijr} - \tilde{d}_{ijr}). \tag{A.2}$$

Related to this error is the within-subject multiple correlation. It indicates how well the configuration fits each subject. If one or more subjects produce a low correlation then these subjects should be removed from the analysis. The square of these correlations can be interpreted as the part of the variance in the dissimilarity data of each subject explained by the configuration. Secondly, an error can be attached to the placement of the stimuli. In the configuration each stimulus is surrounded by a 95 percent confidence region. If all stimuli except one are kept in position, there is a 95 percent probability that the variable stimulus lies in this confidence region.

The average of the within-subject standard errors is called the global standard error. By monitoring this quantity for decreasing dimensionality of the reconstructed psychological space, the proper dimensionality can be found.

A.2 Determining the vector model

Suppose a subject has scaled a single perceptual attribute for all stimuli, producing sensational strengths. Sensational strengths are interval scaled, which means that they are linearly dependent of the perceived strength of the attribute. Each strength should, after a linear transformation, be closely approximated by the value of the perpendicular projection of the stimuli on the attribute direction, which is a vector in the N-dimensional psychological space. By solving the vector model the vector that gives the closest approximation is found.

We will denote the sensational strength of stimulus j for subject r by s_{ir} and the estimated strength, which is given by the projection, by $\tilde{s}_{ir}$. Now the estimated strength is

$$\tilde{s}_{ir} = b_r + \sum_{n=1}^{N} a_{rn} x_{in}. \tag{A.3}$$

Thus the strength is found with a multiple linear regression model, with the estimated strengths as the dependent variables and the coordinates of the stimulus as the independent variables. Through minimizing the error quantity

$$e_r = \sum_{i=1}^{I} (\tilde{s}_{ir} - s_{ir})^2 \tag{A.4}$$

the vector model is solved. The regression coefficients a_{rn} are the subject dependent direction coordinates of the vector.

Now suppose we intend to find a single direction that fits the sensational strengths of all subjects. Therefor we factorize the numbers a_{rn} in two parts, a subject dependent scaling coefficient c_r and a dimension dependent direction coefficient y_n. Now equation (A.3) can be written as

$$\tilde{s}_{ir} = b_r + \sum_{n=1}^{N} c_r y_n x_{in}, \tag{A.5}$$

and the solution of the model is found by minimizing the error

$$e = \sum_{r=1}^{R} \sum_{i=1}^{I} (\tilde{s}_{ir} - s_{ir})^2. \tag{A.6}$$

The coefficients y_n are the subject independent direction coordinates of the vector. The coefficients c_r and b_r determine the linear transformation performed on the projection of the stimulus to approximate the sensational strengths.

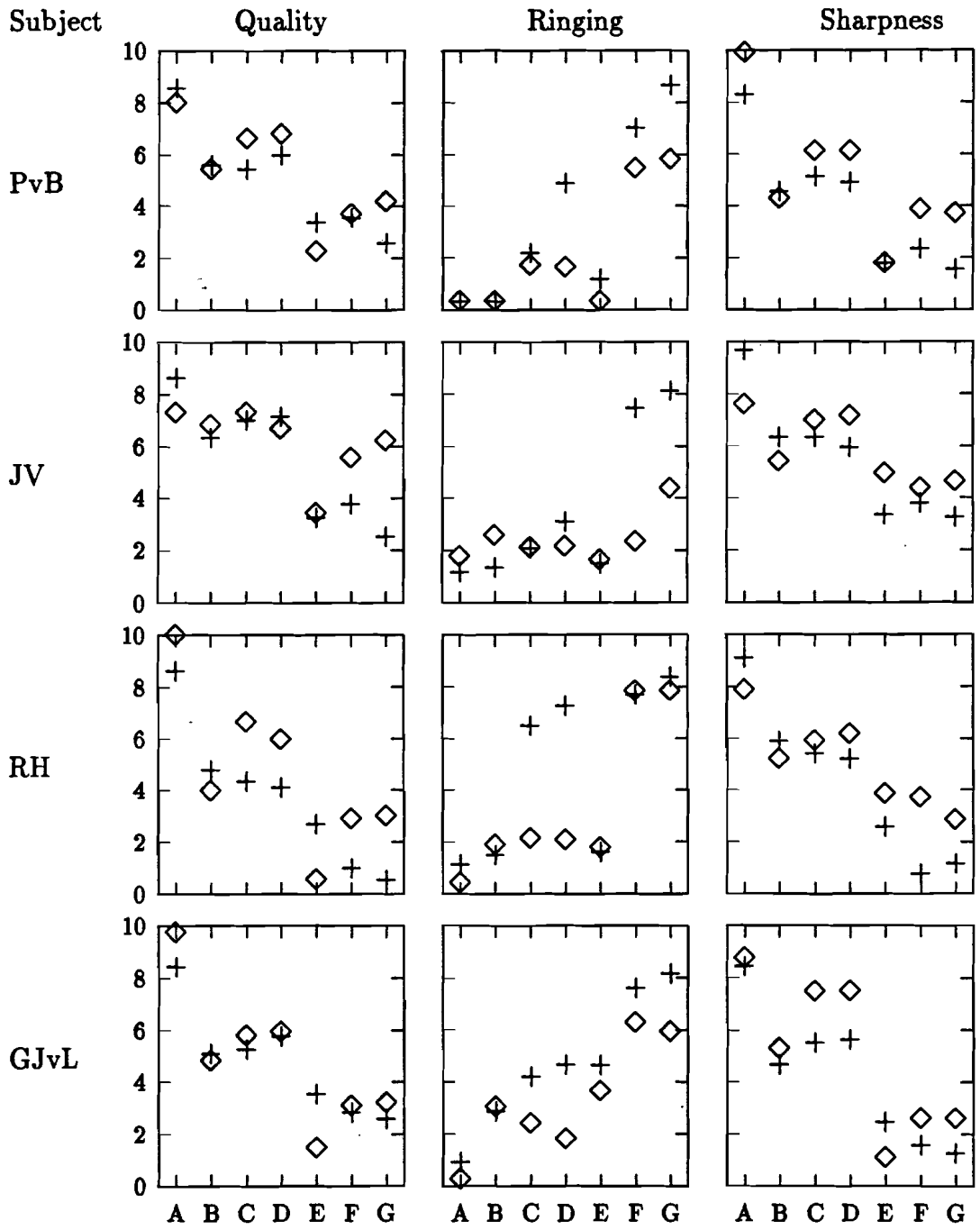
Like the configuration, also the vector found will not fit the sensational strengths data perfectly. Again for each subject a correlation between the strengths and their estimations can be calculated. This is the Pearson correlation, it tells us how well the projection on the vector fits the sensational strengths of each subject.

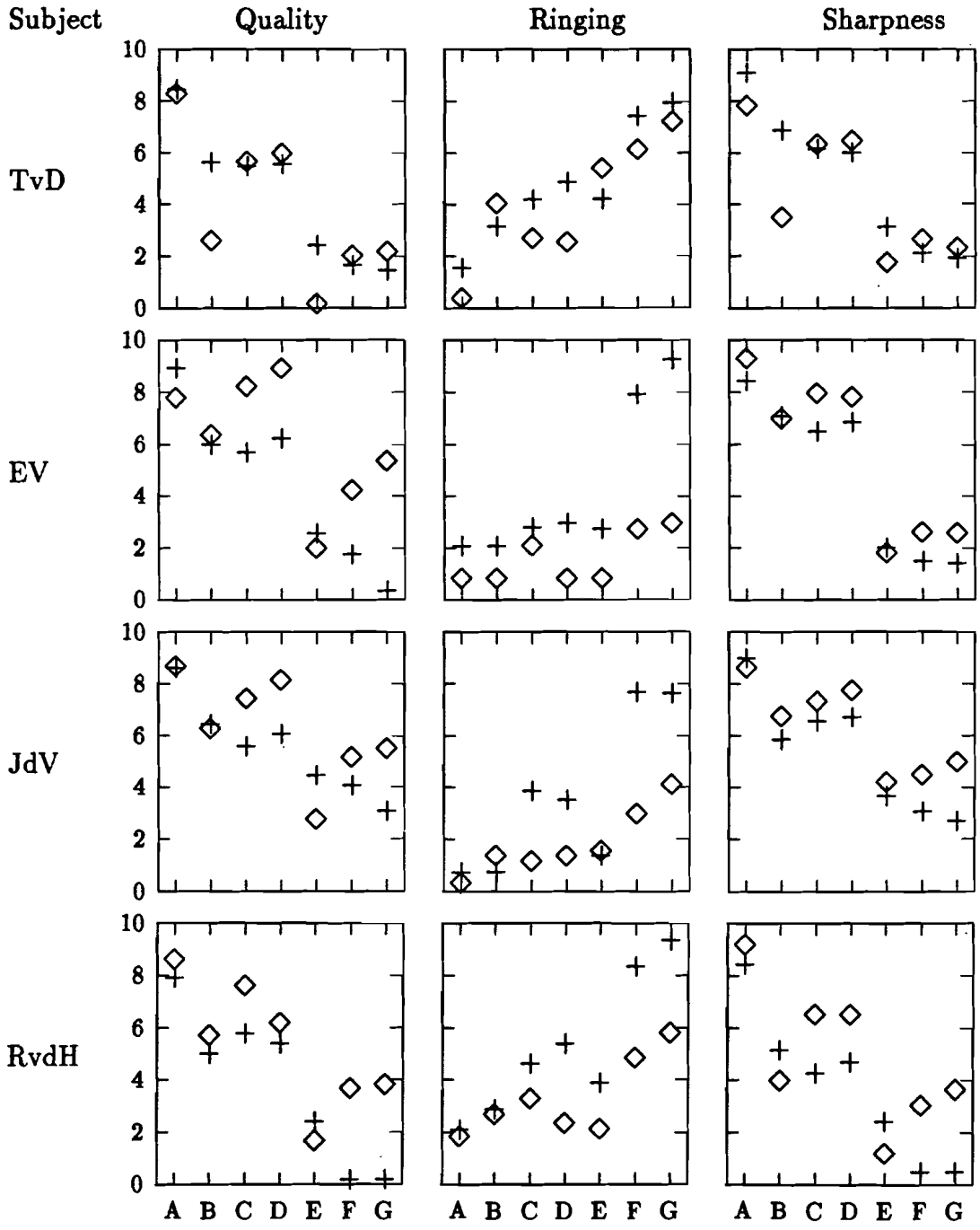
Appendix B

Individual results of the attribute scaling experiments

This appendix contains the figures with the individual results of the attribute scaling experiments done in the multidimensional scaling experiment. For each subject and each attribute scaling a figure is drawn. In this figure the symbol $\diamond$ indicates the results for the scene WANDA and the symbol + the results for the scene STADHUIS. The characters on the horizontal axes indicate the sampling and interpolation filters used. In these B-spline filters the sampling factor m and the order n are varied. In the table below is summarized which combination is meant by which character. On the vertical axes the sensational strengths, determined according to Thurstone's law of categorical judgment, condition D, are depicted.

	A	B	C	D	E	F	G
m	2	3	3	3	4	4	4
n	1	1	5	9	1	5	9


Subject
Quality
Ringing
Sharpness
PvB
JV
RH
GJvL
10
8
6
4
2
0
A B C D E F G

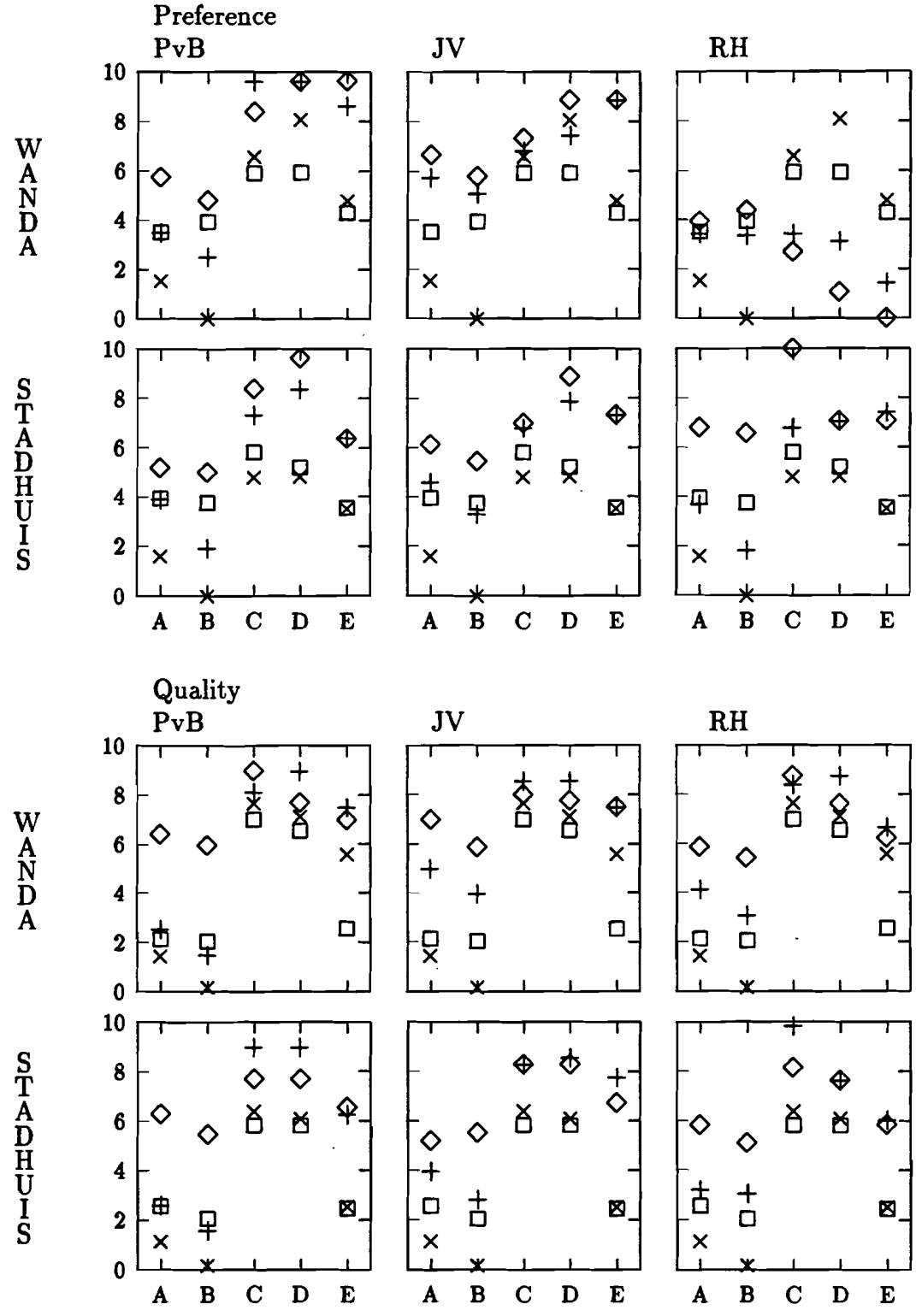

Subject
Quality
Ringing
Sharpness
TvD
EV
JdV
RvdH
0
2
4
6
8
10
A B C D E F G

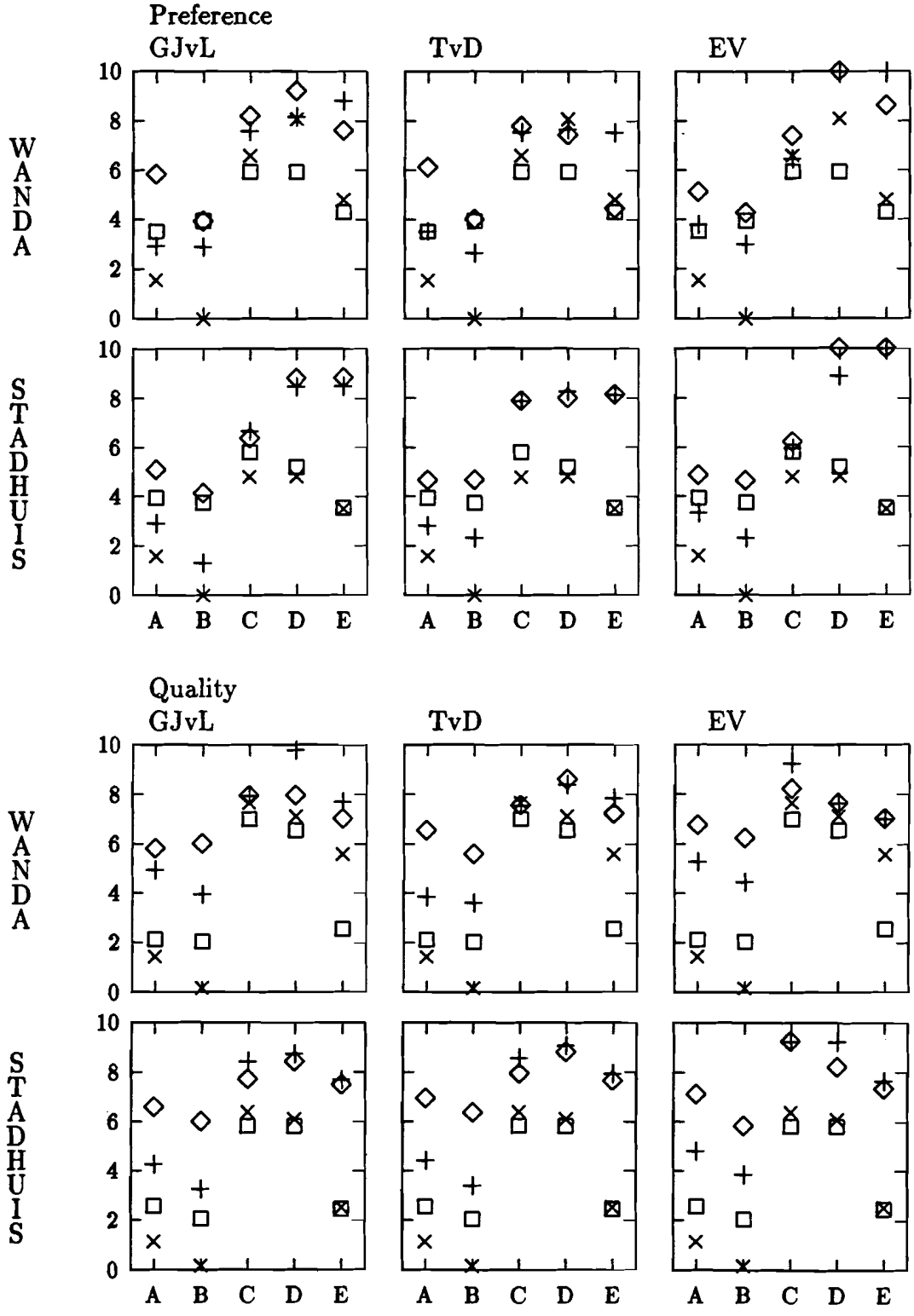
Appendix C

Individual results of the sampled and interpolated images scaling experiments

This appendix contains the figures with the individual results of the quality and preference scaling experiments done with sampled and interpolated images. In the first experiment only the processed image was shown and the subjects gave a preference judgement. In the second experiment the original scene and the processed image were shown next to each other, so the subjects were able to judge the quality of the reproduction. For each experiment, each subject and each scene a figure is drawn. Four different B-spline interpolation filters were used. They are indicated by different symbols in the figures according to the table below. The characters on the horizontal axes indicate the different sampling filters used for each interpolation filter. For A there is no filtering prior to subsampling. B indicates that for sampling the same filter as for interpolation is used. C is sampling with a filter optimized for the used B-spline interpolation filter. And finally, if the interpolation filter is a B-spline of order n, D and E respectively indicate sampling with filters optimized for a B-spline interpolation filter of order $n+1$ and $n+2$. On the vertical axes the sensational strength, determined according to Thurstone's law of categorical judgment, condition D, is depicted.

	$m=2$	$m=3$
$n=1$	◇	□
$n=5$	+	×

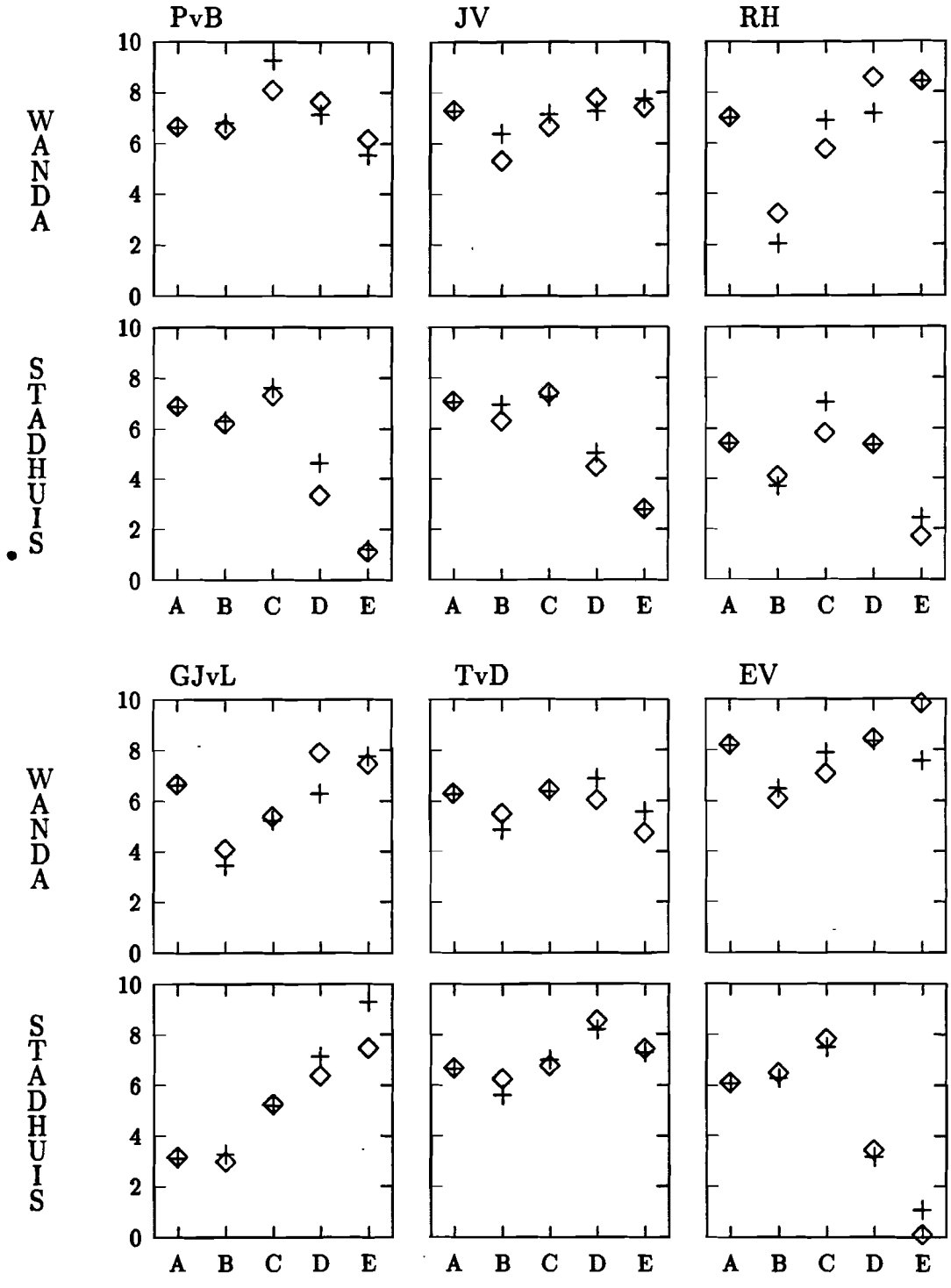

Preference
PvB
JV
RH
WANDA
STADHUIS
Quality
PvB
JV
RH
WANDA
STADHUIS
0
2
4
6
8
10
A B C D E


Preference
GJvL
TvD
EV
WANDA
STADHUIS
Quality
GJvL
TvD
EV
WANDA
STADHUIS
0
2
4
6
8
10
A B C D E

Appendix D

Individual results of the sampled images scaling experiments

This appendix contains the figures with the individual results of the quality scaling experiments done with sampled images that were not interpolated. For each subject and each scene a figure is drawn. In the figures the symbol $\diamond$ indicates the factor two sampled images and the symbol $+$ the factor four sampled images. The characters on the horizontal axes indicate the sampling filters used. For A there is no filtering prior to subsampling. The result is the same for both sampling factors. B, C, D and E respectively indicate the 0, 1st, 2nd and 3rd order B-spline sampling filters. On the vertical axes the sensational strengths, determined according to Thurstone's law of categorical judgment, condition D, are depicted.


PvB
JV
RH
GJvL
TvD
EV
WANDA
STADHUIS
WANDA
STADHUIS
0
2
4
6
8
10
A B C D E