Texture analysis of the radiographic trabecular bone pattern in osteoporosis

J.F. Veenland

Acknowledgments

Texture analysis of the radiographic trabecular bone pattern in osteoporosis

Textuur analyse van trabeculair bot in röntgenfoto's bij osteoporose

Proefschrift

Ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de rector magnificus
Prof. Dr P.W.C. Akkermans M.A.
en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op
woensdag 13 januari 1999 om 15.45 uur

door

Jifke Frouwke Veenland
geboren te Drachten

Promotiecommissie:

Promotores:	Prof. dr E.S. Gelsema
	Prof. dr H.E. Schütte
Overige leden:	Prof. dr ir E. Backer
	Prof. dr ir N. Bom
	Prof. dr P.F. Van der Stelt

Twa Kikkerts

Twa kikkerts oan'e hip dy sprongen om yn't lân.
De ien wie optimist, de oar swier op'e hân.
't Wie melkertsiid, in wein mei folle bussen stûn
Ticht by in hikkedaam. Se sprongen fan 'e grûn
Ynienen rjocht omheech en ploften beide kjel
Troch eigen ûnferstân lyk yn in molkbus del.
Dêr swommen se oerstjoer, de bus wie boardefol,
De glêde kanten lâns, oant ien sei: "k Wit it wol,
Wy bin 'ferlern, sis ik, wy ha it libben hân,
Dit bodzjen jout ús neat, ik brûk leafst myn ferstân.'
Hy bour'le nei omleech, wylst d'oare krigel sei:
'Ik brûk myn poaten leafst en jou mysels net wei'.
Hy sloech yn't wylde wei en tocht:'k Wol net ferhippe.
De molke reamme en skifte en skomme as griene sjippe.
Oant op it lêst de froask nei oerenloang geploeter,
Wol tige wurch fansels, mar libben, siet op bûter.
Doe kaam de boerefeint werom en seach ferheard
Us kikkert yn'e bus en smiet him yn'e feart.

Twee Kikkers

Twee kikkers sprongen rond in het land.
De een was optimist, de ander zwaar op de hand.
't Was melkerstijd, een kar met volle bussen stond
Vlak bij een dam. Ze sprongen van de grond
In een keer recht omhoog en ploften elk
Door eigen onverstand in een bus met melk.
De bus was boordevol, daar zwommen ze overstuur
De gladde kanten langs. De een zei: 'Dit is ons laatste uur
Wij zijn verloren, zeg ik, het leven is voorbij want
Dit ploeteren helpt ons niet, liever gebruik ik mijn verstand'
Hij borrelde naar beneden, terwijl de ander kregel liet horen:
'Ik gebruik liever mijn poten en geef mijzelf niet verloren'.
Hij sloeg in 't wilde weg en dacht:'k Zal niet ten onder gaan.
De melk roomde op en schifte en schuimde als zeep langzaamaan
Totdat op het laatst de kikker, afgemat
Maar levend, na urenlang geploeter op boter zat.
Toen kwam de boerenknecht terug en keek ontzet
Naar onze kikker in de bus en heeft hem in de sloot teruggezet.

Vertaald door H. Veenland, Drachten

Two frogs

Two happy, hopping frogs were having jumping frolic.
One was an optimist, the other melancholic.
't Was milking time, a row of full cans stood aligned
close to their playing field. They jumped, out of their mind,
With one almighty leap and tumbled full of fright,
Together in one can: one left, the other right.
Upset they swam around, their fate increased their fear,
Their limbs slipped down the brim, till one said: 'Oh, my dear,
We've had it now, I say, right here we kick the bucket
I do not fight this mess, my wit says: just go suck it'.
And as he swirly drowned, the other edgy said:
'I will just use my limbs, that will bring me ahead.'
He flapped about and thought while there is life there's hope.
The liquid turned and churned and lathered as soft soap.
Till in the end, the frog, after much splat- and splutter,
Exhausted, yes indeed, but living, sat on butter.
Right then the milk-collector saw this can, from which
He grabbed the lonely frog, to dump him in the ditch.

Vertaald door E.S. Gelsema, Capelle a/d IJssel

Table of contents

Publications and manuscripts based on studies described in this thesis

Veenland JF, Grashuis JL, Gelsema ES. Texture analysis in radiographs: the influence of MTF and noise on the discriminative ability of texture features. Med Phys 1998;25(6):922-936.

Veenland JF, Grashuis JL, van der Meer F, Beckers ALD, Gelsema ES. Estimation of fractal dimension in radiographs. Med Phys 1996;23(4):585-594.

Veenland JF, Grashuis JL, Link TM, Gelsema ES. Reproducibility of texture features in femur radiographs. Proceedings of international workshop "Trabecular bone texture analyis on radiographs", Orleans, December 18-20, 1996.

Veenland JF, Link TM, Konermann W, Meier N, Grashuis JL, Gelsema ES. Unraveling the role of structure and density in determining vertebral bone strength. Calcif Tissue Int 1997;61(6):474-479.

Veenland JF, Grashuis JL, Gelsema ES. Suitability of texture features as indicators of vertebral bone strength. In preparation for publication.

List of published abstracts

Veenland JF, Grashuis JL, Gelsema ES, Beckers ALD, Van Kuijk C. Texture analysis of trabecular bone in radiographs to detect osteoporosis. In: Boehme JM, Rowberg AH, Wolfman NT, ed. Symposium for Computer Assisted Radiology. Winston-Salem: Symposia Foundation, 1994:77-82.

Veenland JF, Gelsema ES, Beckers ALD, van Kuijk C, Grashuis JL. Investigation of bone structure by texture analysis. In: 10th Int. Workshop on bone densitometry. Venice, Italy: Bone and Mineral, 1994;25(Suppl. 2)S12.

Veenland JF, Grashuis JL, Gelsema ES, van Kuijk C. The value of fractal analysis for characterization of trabecular structures in hand radiographs. In: European Congress of Radiology. Vienna: Eur Radiol, 1995;5:S174-175.

Veenland JF, Link TM, Konermann W, Meier N, Grashuis JL, Gelsema ES. Can texture parameters predict bone strength? In: 11th Int. Workshop on bone densitometry. Gleneden Beach: Osteoporosis International, 1996;6:91.

Veenland JF, Grashuis JL, Beckers ALD, Gelsema ES. The dependence of fractal dimension estimates on grey-level variance. In: 11th Int. Workshop on bone densitometry. Gleneden Beach: Osteoporosis International, 1996;6:104.

Veenland JF, Link TM, Konermann W, Meier N, Grashuis JL, Gelsema ES. Comparison of two fractal dimension estimation methods in predicting fracture risk in vertebrae [poster abstract PSu 253]. In: 1996 World Congress on Osteoporosis. Amsterdam: Osteoporosis Int., 1996;6(Suppl 1):S146.

Veenland JF, Link TM, Konermann W, Meier N, Grashuis JL, Gelsema ES. Sex differences in predictive power of vertebral strength of texture parameters (poster). In: Fifth Bath Conference on Osteoporosis and Bone Mineral Measurements. British Institute of Radiology, London, 1996;27-28.

Chapter 1

General introduction

Texture

Texture is an image property which is difficult to grasp. It can be described as a "homogeneous visual pattern"[1]. but there exists no formal definition of texture. Intuitively people can discriminate between different textures, referring to visual clues like coarseness, orientation, periodicity, and regularity. Using such concepts, several authors have tried to quantify these aspects of texture[2]. However, texture encompasses more than these more or less random aspects to which the human eye is sensitive. Therefore, the majority of texture analysis algorithms is based on an image model, in which certain characteristics of the image texture are condensed. Using this image model, texture features can be derived, most of which cannot be related to visual image features.

Texture analysis methods are able, in contrast to a human observer, to quantify textures objectively. Therefore, texture features can be used for the purpose of characterization, discrimination, and segmentation of textures in, for example, aerial and satellite imagery. Most texture analysis methods have been developed and tested on textures from the collection of texture images in Brodatz[3] before putting them into use in a more realistic environment. Since the early seventies, texture analysis methods have also been applied in medical images. For example, Sutton et al. tried to categorize different stages of pulmonary disease in radiographs[4]. Since then, the field of application of texture analysis methods in radiology has expanded from chest radiographs to mammograms and bone radiographs.

The goal of our study is twofold: in the first place to assess the suitability of different texture analysis methods for use in radiographs, secondly to select or develop texture features which are able to quantify the changes in the radiographic trabecular pattern occurring in osteoporosis.

Osteoporosis

Osteoporosis is defined as "a disease characterized by low bone mass and microarchitectural changes of bone tissue, leading to enhanced bone fragility and a consequent increase in fracture risk." (WHO, 1994)[5]. Until recently, the bone mineral density was regarded as the main characteristic of osteoporosis. Several sophisticated techniques have been developed to measure density, for example Quantitative Computed Tomography (QCT) and Dual-energy X-ray Absorptiometry (DXA). More recently, ultrasound techniques have gained terrain

in diagnosing osteoporosis. However, osteoporosis in not only characterized by a decrease in density, but also by structural changes in the architecture of bone. In this study we have focussed on the structural changes in the trabecular bone. The osteoporotic changes occurring in different bones have been described using techniques from histomorphometry. A less invasive procedure is offered by radiographic imaging: the 3-dimensional structure is projected in the radiograph, where it appears as texture. Several authors have described the osteoporotic changes in the radiographic trabecular texture. In vertebrae, it is noted that the "relative accentuation of the vertical trabeculae leads to vertical radiodense striations"[6]. "There is preferential resorption of transverse trabeculae and accentuation of the remaining trabeculae"[7]. A semiquantitative approach is followed by Singh et al.[8]. According to the appearance of the trabecular bone in different areas in the proximal femur, the pattern is graded into one of seven groups: the Singh-index. Cockshott tried to establish a Calcaneal Morphologic Index for the os calcis, but was hampered by a poor observer reproducibility[9]. The vertebral trabecular pattern has also been indexed into four categories[10]. These indices have a number of drawbacks. In the first place, the intra- and inter observer variability tends to be high. Second, since the categorization is rather coarse, only large differences can be detected. Finally, in these indices radiographic density and structure are combined: the contribution of structure and density cannot be separated. Since the radiographic density is not only determined by the bone mineral density, but also by the amount of overlying tissue and the settings of the radiographic equipment, the categorization may vary with these factors. Since it is not possible for the human eye to assess structure independent of the density present in the image, computer analysis is inevitable. The use of computerized texture analysis methods for the quantification of the radiographic trabecular pattern has several advantages. In the first place, the structure can be evaluated independently of the radiographic density. Secondly, with a computerized method a more objective and reproducible result can be realized than with human visual inspection. In the third place, the structure can be evaluated using continuous scales.

Requirements for texture features

As mentioned above, the goal of our study is to find or develop texture features which are able to quantify the changes in the radiographic trabecular pattern occurring in osteoporosis. In order for a texture feature to suit this objective, a

number of requirements should be fulfilled. In the following we will describe and amplify these requirements.

Each digital image is composed of pixels with varying gray-levels. The gray-level histogram, specifying for each gray level the number of occurrences in the image, gives no information about the spatial distribution of the gray-levels. Therefore, the histogram contains no texture information. Whereas one image is related to only one histogram, the inverse is not true: images with different textures can be characterized by the same histogram. On the other hand, texture features quantify the spatial pattern laid out by the different gray-levels in the image. In literature, a wide variety of texture analysis methods has been described. All these methods try to quantify the local spatial distribution of gray-levels using different image models. For example, the Spatial Gray-Level Dependence Method (SGLDM) quantifies the occurrence of pairs of gray-levels at certain distances, and under given angles, whereas in the Fourier Method the image is regarded as a set of spatial frequencies each with a certain amplitude and phase.

Since texture information is hidden in the coherence of gray-levels, and not in the absolute values of the occurring gray-levels, texture features should be able to characterize texture information independent of the absolute gray-level values. This can be formulated as the following requirement

Texture features should be able to quantify "real texture"; this means that the features should be invariant under strict monotonic gray-level transformations (1)

This criterion was already formulated by Haralick[11]. According to Haralick, texture features should be invariant under monotonic gray-level transformations introduced by variations in lighting, lens, film, developer, and digitizer. A less strict criterion can be used when the shapes of the histograms are similar: texture features should be invariant under linear transformations of the gray level histogram. In that case, the texture features fulfil the most basic criterion: independence of the mean and variance of the gray-level histogram. None of the methods described in literature is inherently invariant under monotonic gray level transformations. This property can be obtained, at the expense of some information loss, by a preprocessing step like thresholding or histogram equalization. In our study we decided not to use methods which work only on thresholded images. We feel that by the thresholding step too much information, hidden in the gray-level transitions is lost. On the other hand, we accepted histogram equalization as a preprocessing step, since, compared to thresholding, histogram equalization preserves more gray-level transition information.

5

We focus in our study on four different texture analysis methods. Two methods, the Spatial Gray-level Dependence Method and the Fourier Method have been selected from texture analysis literature, based on their superior discriminative performance on aerial photographs[12]. Since the fractal dimension has recently become a popular feature for quantifying texture in radiographs, different fractal dimension estimation methods are combined to form the third method. The fourth method, based on morphological gradients, has been developed by us, especially for the purpose of quantifying trabecular structure in bone radiographs.

Texture analysis methods can be used for different purposes such as characterization, discrimination and segmentation of textures. Although these goals differ, the basic requirement is the same: a texture analysis method should be able to extract that information from an image, which uniquely describes it. Clearly, a texture feature which cannot discriminate between two different textures, is not very useful for characterization or segmentation purposes. Therefore, we can formulate a second basic requirement for texture features:

Texture features should be able to discriminate between different textures   *(2)*

Texture analysis methods have been applied to all kinds of imagery, ranging from aerial and satellite images to medical images. In this study, we focus on the application of texture analysis methods to radiographs. Radiographs differ essentially from other imagery in the sense that they are dose limited: for patient safety, the exposure dose is kept as low as possible. This has two degrading effects on image quality. In the first place, the quantum mottle. Quantum mottle, the most important source of noise in radiographs, is the result of statistical fluctuations in the number of X-ray quanta per unit area absorbed by the intensifying screen. In the second place, the intensifying screens which are used in front of the film in order to enhance the light output, have a blurring effect. The photons, generated in the intensifying screen by the absorption of X-ray quanta, are scattered in the screen before being absorbed by the film. Due to this scattering process, a point in the object is depicted as a blob in the image. This blurring effect results in a reduced spatial resolution, contrast and sharpness. In the spatial frequency domain, this scattering effect is described by the Modulation Transfer Function (MTF). The MTF quantifies the relative transmittance of the spatial frequencies occurring in the depicted object.

Consequently, noise and image blur are both inherently present in radiographs. In order to be able to discriminate between different textures in radiographs, the following requirement should be fulfilled:

The discriminative power of texture features should be relatively independent of noise and the modulation transfer function (3)

In radiography, a 3-dimensional object is projected on a 2-dimensional plane. When the object is non-homogenous, as is the case with trabecular bone, rotations will have a profound effect on the structure visible in the radiograph. Whereas a texture analysis method may be expected to detect changes in texture due to rotation over large angles, it should be robust for texture variations caused by small rotations. In a clinical setting, such small rotations due to variations in positioning are inevitable. Furthermore, no two radiographs of one object will be exactly the same, even if the object is not repositioned. In each of the different components of the imaging chain small random variations may occur. While large variations, due to errors, must be detected, the texture features should be robust within the limits of reproducibility of experimental conditions. This can be formulated in the following requirement:

Texture features should be reproducible under repositioning, including small rotations (4)

The requirements as listed above, are applicable for texture analysis in general and texture analysis in radiographs. In the following, a number of requirements will be discussed which are tailored to the goal of our study: to select or develop texture features which are able to quantify the radiographic trabecular pattern in osteoporosis.

In the first part of the definition of osteoporosis as "a disease characterized by low bone mass, microarchitectural changes of bone tissue leading to enhanced bone fragility ...." the relation between BMD, structure and strength is outlined. Assuming that texture features are able to quantify the structural status of trabecular bone, these features should be related to the strength of trabecular bone. Using a skeletal site where the bone is primarily trabecular, the following requirement can be formulated:

Texture features should correlate with bone strength (5)

Let us assume that the texture features used fulfil requirement 1. In that case, the correlation with bone strength cannot be attributed to differences in radiographic film density. However, the texture features may still be dependent on the bone mineral density (BMD). The intriguing problem is that structure and density are two concepts which are physically closely intertwined and therefore strongly correlated: no structure without density and no density without structure. Therefore, each structure feature will show a correlation with density. The question

then is which part of the discriminative power can be attributed to BMD, and which part to structural differences. By using a separate index for BMD, these effects can be isolated. A more restrictive requirement can be formulated:

Texture features should correlate with strength, independent of bone mineral density (5a)

Since osteoporosis is a systemic condition, we can assume that the structural status of trabecular bone at different skeletal sites is correlated. When, for example, the diagnosis osteoporosis is based on the presence of vertebral fractures, not only the trabecular structure in the vertebrae, but also the structure in, e.g., hip, wrist, and hand radiographs will show osteoporotic changes. Radiographs of these, non-fractured, sites can be used to study whether osteoporotic changes occurring in trabecular bone can be detected by texture analysis.

Texture features should be able to discriminate between osteoporotic and non-osteoporotic trabecular bone (6)

Again assuming that requirement 1 holds, the differences found in texture between osteoporotic and healthy trabecular bone cannot be attributed to differences in radiographic film density. However, analogous to the reasoning above, the texture features may still be dependent on the bone mineral density (BMD). Therefore, the requirement should be formulated more restrictively:

Texture features should be able to discriminate between osteoporotic and non-osteoporotic trabecular bone, independent of bone mineral density (6a)

The requirements listed above are elaborated in Chapters 2 through 8.

The methods used for texture analysis throughout this study are described in detail in Chapter 2. Moreover, their respective invariance under gray-level transformations (requirement 1), is discussed.

In Chapter 3, we compare the discriminative power of the different texture analysis methods, described in Chapter 2 (requirement 2). For that purpose we used two series of textures: fractal and Brodatz. Whereas the Brodatz images contain miscellaneous textures, the differences between the fractal images are more gradual. The discriminative performance of the individual features is compared, as is the performance when a number of features of each method are combined. In addition, since we are interested in the performance of the different methods in circumstances mimicking the conditions prevailing in radiographs, the effects of noise and image blur, as the main causes of degradation in the radiographic process,

are simulated on both series of textures. This enables us to evaluate the influence of both factors on the discriminative performance of, the individual features of, the texture analysis methods (requirement 3). Which method performs best on which image type? How do noise and image blur affect the different texture analysis features? Which method is the least sensitive to noise and image blur? These questions will be answered in Chapter 3.

The effect of noise and image blur on fractal dimension estimation methods is studied in more detail in Chapter 4 (requirement 3). Since images with a known fractal dimension can be generated, these images can be used as a gold standard for the fractal dimension estimation methods. Apart from comparing the outcomes of the different methods, these images are used as the basis for the simulation of the effect of noise and image blur. In this way, the influence of noise and image blur on the dimension estimation is studied directly. In what way are the different fractal dimension estimation methods influenced by noise and image blur? What is the influence on the estimated fractal dimensions?

Chapter 5 contains an elaboration of requirement 4. The reproducibility of texture parameters is an often neglected but nevertheless important issue. Reproducibility is defined as the degree of consistency among repeated measurements taken on a given object under conditions that are as close to identical as experimentally possible. For a texture analysis method to be clinically useful, it should be reproducible under repositioning. It should be robust for small texture variations caused by repositioning which inherently occur in a clinical setting. In Chapter 5, we study the reproducibility of the texture features of the four texture analysis methods applied to the trabecular pattern of femora.

The correlation of texture features with strength (requirement 5) is studied in Chapter 6. In this study, we use vertebrae for two reasons. In the first place, vertebrae consist mainly of trabecular bone. Secondly, the simulation of the biomechanical fracture conditions is not as complicated as for other osteoporotic fracture prone skeletal sites. In this study, texture features and BMD are determined before each vertebra was crushed. Since the strength of the vertebra is not only determined by the architecture of the bone, but also by the size of the vertebra, the force necessary to fracture the vertebra is corrected for the vertebral size resulting in the parameter fracture stress. The correlation coefficients of the texture features with fracture stress are determined. By using partial correlations, we are able to separately evaluate the contribution of structure and density to vertebral strength.

9

In Chapter 7 and 8, we investigate the ability of texture features to discriminate between osteoporotic and healthy trabecular bone (requirement 6). In Chapter 7 a region of interest with a visually homogenous trabecular structure in the proximal phalanx of the right index finger is used. In the same region and on the same radiograph BMD is determined, using a method described by Trouerbach et al.[13]. By performing a matching, based on BMD, of the regions of interest of osteoporotic and non-osteoporotic bone, we are able to study the differences in texture in osteoporotic and healthy bone, independent of density. In Chapter 8, the hand radiographs of 118 females with one or more vertebral deformities, and of 87 females with no vertebral deformities are used to investigate the suitability of different texture features for detecting osteoporotic changes in the trabecular bone structure.

In Chapter 9, the topics investigated in the foregoing chapters are discussed and mutually related. Chapter 10 contains the summary.

References

1. Tomita F, Tsuji S. Computer analysis of visual textures. Boston: Kluwer Academic Publishers, 1990.
2. Tamura H, Mori S, Yamawaki T. Textural features corresponding to visual perception. IEEE Trans Syst Man Cybern 1978;8(6):460-473.
3. Brodatz P. Textures, a photographic album for artists & designers. (1st ed.) New York: Dover Publications, 1966.
4. Sutton RN, Hall EL. Texture measures for automatic classification of pulmonary disease. IEEE Trans Comp 1972;C-21(7):667-676.
5. WHO studygroup. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. WHO 1994.
6. Resnick D, Niwayama G. Diagnosis of bone and joint disorders. (2nd ed.) Saunders Company, 1988.
7. Greenfield GB. Radiology of bone diseases. (4 ed.) Philadelphia: J.B. Lippincot Company, 1986.
8. Singh M, Riggs BL, Beabout JW, Jowsey J. Femoral trabecular pattern index for evaluation of spinal osteoporosis. Mayo Clin Proc 1973;48:184-189.
9. Cockshott WP, Occleshaw FRCP, Webber C, Walter SD, O'Brien K. Can a calcaneal morphologic index determine the degree of osteoporosis. Skeletal Radiology 1984;12:119-122.
10. Schnitzler CM, Pitchford DGK, Willis EM, Gear KA. Comparison of the radiographic vertebral trabecular pattern with the vertebral fracture prevalence and spinal bone density. Osteoporosis Int 1993;3:293-299.

11. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern 1973;3(6):610-621.
12. Weszka JS, Dyer CR, Rosenfeld A. A comparative study of texture measures for terrain classification. IEEE Trans Syst Man Cybern 1976;6:269-285.
13 Trouerbach WT, Birkenhäger JC, Collette BJA, Drogendijk AC, Schmitz PIM, Zwamborn AW. A study on the phalanx bone mineral content in 273 normal pre- and post-menopausal females (transverse study of age-dependent bone loss). Bone and Mineral 1987;3:53-62.

Chapter 2

Texture analysis methods

1. Introduction

In our investigations of the applicability of texture analysis methods to radiographs, and, more specifically, the suitability of these methods for the quantification of osteoporosis, we focused on four different texture analysis methods. Two of these methods, the Spatial Gray-Level Dependence Method and the Fourier Power Spectrum method were selected from texture analysis literature, based on their superior discriminative performance on aerial photographs[1]. Since the fractal dimension has recently become a popular feature for quantifying texture in radiographs, different fractal dimension estimation methods were combined to form the third method. The fourth method, based on morphological gradients, was developed by us, especially for the purpose of quantifying trabecular structure in bone radiographs. In this chapter, the four methods will be described extensively. For each of the methods, the invariance under strict monotonic (requirement 1 in the introduction), respectively linear gray-scale transformations will be discussed.

2. Spatial Gray-Level Dependence Method

In 1973 Haralick et al.[2] introduced the spatial gray-level dependence matrix, also called the co-occurrence matrix, and its derived features. The method consists of three steps. First, a histogram equalization is performed: the number of gray-levels in the image is reduced to a fixed number G and the gray-levels are redistributed in such a way that the probability of occurrence for all gray-values is equalized. Second, a matrix $M_{d,\theta}$ is composed, the elements $m_{d,\theta}(i,j)$ of which are the co-occurrences of gray-levels i and j in the image over a distance d, under an angle θ. Normalizing the matrix elements for the total number of co-occurrences yields a probability matrix $p_{d,\theta}(i,j)$. From this matrix, as a third step, a variety of features can be computed. We used five features, defined by Haralick et al. as Angular Second Moment, Contrast, Correlation, First Diagonal Moment, and Entropy:

$$ASM_{d,\theta} = \sum_{i,j=1}^{G,G} p_{d,\theta}^2(i,j) \tag{1}$$

$$CON_{d,\theta} = \sum_{i,j=1}^{G,G} (i-j)^2 p_{d,\theta}(i,j) \tag{2}$$

15

$$COR_{d,\theta} = \frac{\sum_{i,j=1}^{G,G} (i - \mu_i)(j - \mu_j) \, p_{d,\theta}(i, j)}{\sigma_i \sigma_j} \tag{3}$$

$$\mu_x = \sum_{x=1}^{G} x \quad \sigma_x^2 = \sum_{x=1}^{G} (x - \mu_x)^2 \quad with \ x = i, j$$

$$DMO \, 1_{d,\theta} = \sum_{i,j=1}^{G,G} \sqrt{\frac{|i - j| \, p_{d,\theta}(i, j)}{2}} \tag{4}$$

$$ENT_{d,\theta} = \sum_{i,j=1}^{G,G} p_{d,\theta}(i, j) \log(p_{d,\theta}(i, j)) \tag{5}$$

We will refer to the features generated by this method as the co-occurrence features.

Invariance under strict monotonic increasing gray-level transformation
Histogram equalization, or histogram linearization, performs a transformation on the gray-levels occuring in an image with the effect that the resulting gray-levels have a uniform density distribution[3]. The transformation function (*T*) used is the cumulative distribution function:

$$T(k) = \sum_{j=0}^{k} p(j) \tag{6}$$
$$k = 0, 1, .., L - 1,$$

where *p(j)* is the probability of occurrence of the *j*th gray level, and *L* is the number of gray-levels.

A transformation of the gray-level *k* will, as long as this transformation is strictly monotonic increasing, not influence the probability density distribution *p*, and the transformation *T*. Therefore, histogram equalization of a gray-level transformed image will result in the same "equalized" image as the equalization of the original image. This only holds as the transformation is strictly monotonic increasing.

Since the histogram equalization step neutralizes the effect of strict monotonic increasing gray-level transformations of the image, the co-occurrence features

which are extracted from the equalized image, are invariant under such transformations.

3. Fourier Method

In the Fourier Method we merge three sets of features which have as a common denominator that they can all be derived from the Fourier Power Spectrum. The first group encompasses the oldest and most widely used spectrum features described by Bajcsy[4], quantifying the radial and angular distribution of the Power Spectrum $P(u,v)$:

$$P_{r_1,r_2} = \sum_{r_1^2 \leq u^2 + v^2 < r_2^2} P(u,v) \tag{7}$$

$$P_{\theta_1,\theta_2} = \sum_{\theta_1 \leq \tan^{-1}(v/u) < \theta_2} P(u,v). \tag{8}$$

The second group consists of two features introduced by Katsuragawa et al.[5]: the root-mean-square variations (*RMS*), and the first moment (*FMO*), which are defined as follows:

$$RMS = \sqrt{\sum_{u,v} P(u,v)} \tag{9}$$

$$FMO = \frac{\sum_{u,v} \sqrt{u^2 + v^2}\, P(u,v)}{\sum_{u,v} P(u,v)} \tag{10}$$

While the first feature measures the square root of the total power present, the second feature expresses the mean power weighed with the frequency.

The features of the third group are described by Liu et al.[6]. This group of features contains both "local" spectrum features, such as the amount of energy in the major peak, and "global" spectrum features, such as the isotropy of the power spectrum. The definition of the features used in this study are listed in Table 1. All the features are based on the normalized spectrum $p(u,v)$.

Table 1. *Power Spectrum features described by Liu[6]*

Energy in major peak

$$f_1 = p(u_1, v_1) \tag{12}$$

Laplacian of major peak

$$f_2 = \nabla^2 P(u_1, v_1) = P(u_1 + 1, v_1) + P(u_1 - 1, v_1) \\ + P(u_1, v_1 + 1) + P(u_1, v_1 - 1) - 4P(u_1, v_1) \tag{13}$$

Laplacian of secondary peak

$$f_3 = \nabla^2 P(u_2, v_2) \tag{14}$$

Squared major peak frequency

$$f_5 = u_1^2 + v_1^2 \tag{15}$$

Relative orientation of major and secondary peaks

$$f_6 = \left| \tan^{-1}(\frac{v_1}{u_1}) - \tan^{-1}(\frac{v_2}{u_2}) \right| \tag{16}$$

Major peak horizontal frequency

$$f_9 = u_1 \tag{17}$$

Major peak vertical frequency

$$f_{10} = v_1 \tag{18}$$

Secondary peak horizontal frequency

$$f_{11} = u_2 \tag{19}$$

Secondary peak vertical frequency

$$f_{12} = v_2 \tag{20}$$

Squared distance between major and secondary peak

$$f_{13} = (u_1 - u_2)^2 + (v_1 - v_2)^2 \tag{21}$$

Isotropy of power spectrum

$$f_7 = \frac{|\sigma_u - \sigma_v|}{\sqrt{(\sigma_u + \sigma_v)^2 - 4\sigma_{uv}^2}} \tag{22}$$

Where

$$\sigma_u = \sum_u \sum_v u^2 p(u, v), \sigma_v = \sum_u \sum_v v^2 p(u, v), \\ \sigma_{uv} = \sum_u \sum_v uv p(u, v) \tag{23}$$

Moment of inertia, quadrant I

$$f_{17} = \sum_{u>0} \sum_{v>0} \sqrt{u^2 + v^2} \, p(u, v) \tag{24}$$

Moment of inertia, quadrant II

$$f_{18} = \sum_{u<0} \sum_{v>0} \sqrt{u^2 + v^2}\, p(u,v) \tag{25}$$

Moment ratio

$$f_{19} = \frac{f_{18}}{f_{17}} \tag{26}$$

Percentage energy, quadrant I

$$f_{20} = \sum_{u>0} \sum_{v>0} p(u,v) \tag{27}$$

Percentage energy, quadrant II

$$f_{21} = \sum_{u<0} \sum_{v>0} p(u,v) \tag{28}$$

Relative entropy of power spectrum, R_1

$$f_{25} = \frac{-\sum\limits_{u,v \in R_1} p_1(u,v) \log p_1(u,v)}{K_1} \tag{29}$$

where

$$p_i(u,v) = \frac{P(u,v)}{\sum\limits_{u,v \in R_i} P(u,v)} \tag{30}$$

K_i = number of distinct frequencies in R_i,

$R_i = u, v$, such that

$$\frac{i-1}{4} u_m < |u| < \frac{i}{4} u_m, \text{ and } \frac{i-1}{4} v_m < |v| < \frac{i}{4} v_m$$

u_m, v_m are maximum frequency components for the spectrum.

Relative entropy, R_2

$$f_{26} = \frac{-\sum\limits_{u,v \in R_2} p_2(u,v) \log p_2(u,v)}{K_2} \tag{31}$$

Relative entropy, R_3

$$f_{27} = \frac{-\sum\limits_{u,v \in R_3} p_3(u,v) \log p_3(u,v)}{K_3} \tag{32}$$

Relative entropy, R_4

$$f_{28} = \frac{-\sum\limits_{u,v \in R_4} p_4(u,v) \log p_4(u,v)}{K_4} \tag{33}$$

When $P(u,v)$ is the power spectrum, the normalized spectrum is given by

$$p(u,v) = \frac{P(u,v)}{\sum\limits_{u,v \neq 0} P(u,v)} \tag{11}$$

In the following text, we will refer to the whole set of features, which can be derived from the Fourier Power Spectrum, as the Fourier features.

Invariance under strict monotonic increasing gray-level transformation
The first and second group of features are computed from a non-normalized spectrum. Therefore, these features are not invariant under linear gray-level transformations, let alone monotonic increasing gray-level transformations. Of the third group, all features except for the ones which compute the Laplacian of the primary and secondary peak (f_2 and f_3) are invariant under linear gray-scale transformation. The features which are based on frequencies of the spectrum (f_5, f_9 through f_{13}) are even invariant under strict monotonic increasing gray-level transformations.

4. Fractal Dimension

In literature, a wide variety of fractal dimension estimation methods has been described. These methods are all based on the same principle: an image characteristic is measured as a function of a scale parameter. For fractal images, the relation between these two quantities can be described by a straight line in a log-log domain. The slope, computed by linear regression analysis, is linearly related to the fractal dimension. The various estimation methods differ in the definition of the image characteristic and of the scale parameter. For our study, we selected two methods which have been applied to radiographic images.

A. Power Spectrum Method
An image intensity surface can be modelled by a two-dimensional fractal Brownian function[7]. The Fourier power spectrum $P(f)$ of a fractal Brownian function is related to the spatial frequency f by the following equation

$$P(f) \propto f^{-(2H+1)} \tag{34}$$

In the Power Spectrum Method, the Hausdorff-Besicovitch dimension H is computed from the slope of the straight line fitted in the log $P(f)$ versus log f plot. The Hausdorff-Besicovitch dimension is related to the fractal dimension D by

$$D = E + 1 - H \qquad (35)$$

where E is the topological dimension of the surface. Since we reduced the two-dimensional power spectrum of each image to a one-dimensional spectrum by averaging over circles with radius f, the topological dimension is equal to one .

B. Blanket Method

Peleg et al.[8] extended the Blanket Method, described by Mandelbrot[9], to three dimensions. The surface area of an object can be estimated by measuring the volume between an upper blanket $u_\varepsilon(i,j)$ and a lower blanket $b_\varepsilon(i,j)$ which are not further than a distance ε above and below the surface to be measured. The upper and lower blanket are defined, respectively, as follows:

$$u_\varepsilon(i,j) = MAX\{ u_{\varepsilon-1}(i,j) + 1, \underset{|(m,n)-(i,j)|\leq 1}{MAX} u_{\varepsilon-1}(m,n) \} \qquad (36)$$

$$b_\varepsilon(i,j) = MIN\{ b_{\varepsilon-1}(i,j) - 1, \underset{|(m,n)-(i,j)|\leq 1}{MIN} b_{\varepsilon-1}(m,n) \} \qquad (37)$$

where (m,n) are the four-connected neighbors of (i,j). Initially (for $\varepsilon=0$), both blankets coincide with the intensity surface of the image. For each ε, the volume V between the blankets can be computed using

$$V(\varepsilon) = \sum_{i,j} \{ u_\varepsilon(i,j) - b_\varepsilon(i,j) \} \qquad (38)$$

Peleg et al.[8] define the area of the surface $A(\varepsilon)$ as half of the volume increase:

$$A(\varepsilon) = \frac{1}{2} \{ V(\varepsilon) - V(\varepsilon - 1) \} \qquad (39)$$

Using the relation

$$A(\varepsilon) \propto \varepsilon^{2-D} \qquad (40)$$

the fractal dimension of the surface can be determined from the log $A(\varepsilon)$ versus log ε plot.

In this method the unit of measurement in the intensity direction of the image and the unit of measurement in the (x,y) plane are treated in the same way, as is demonstrated by the use of the "+1" and "−1" term in the respective definitions of the upper and lower blanket. The result of this erroneously assumed equivalence is a dimension estimate which varies when the intensity scale is changed independently of the scale in the (x,y) plane, thus violating the scale independence requirement. Lynch et al.[10] met this objection by omitting the "+1" and "−1" terms in equation 36 and 37, respectively. The max and min operators thus defined correspond with the grey-scale dilation and erosion from mathematical morphology. In the following Chapters we will use this modified version of the Blanket Method.

Linear Regression Analysis
The fractal dimension is computed by fitting a regression line in a plot of the log of the image characteristic versus the log of the scale parameter. It is well recognized that the fractal behaviour of natural images is limited to a certain range of the scale parameter. Therefore, the linear regression analysis should not take all data-points, deduced from all available resolution scales, into account: an upper and lower boundary must be specified for the scale parameter. In the following Chapters, we will specify for each study the boundaries chosen.

Invariance under strict monotonic increasing gray-level transformation
Both fractal dimension estimation methods are not invariant under strict monotonic gray-level transformations, but they are both invariant under linear gray-level transformations. For the Blanket Method, the invariance under addition of a constant gray-level, is effected by the definition of the volume V as the difference between the two blankets. The invariance under addition of a constant gray-level on the fractal dimension estimated by the Power Spectrum Method is explained by the fact that except for a change of power in the frequency origin, the addition of a constant to the image has no influence on the power spectrum. For both methods applies that the invariance under a global multiplicative gray-level transformation is induced by the proportionality relation which forms the heart of both methods (eq. 34 and eq. 40, respectively).

5. Morphological Gradient Method

The theory of Mathematical Morphology is described extensively by Serra[11]. Mathematical Morphology provides, inter alia, techniques for the extraction of image features. Although the basic operations are originally defined for binary images, they can easily be extended for gray-tone images. Each operation is defined with a structuring element, which determines the size, shape and orientation of the local neighborhood on which the operation is performed. In gray-scale images the two basic operations, erosion and dilation, are defined using minimum and maximum filters with the neighborhood determined by the structuring element. If $f(x,y)$ is the input image and $b(x,y)$ the structuring element, erosion $\ominus$ and dilation $\oplus$ are defined, respectively as[3]

$$(f \ominus b) = MIN\{f(s + x, t + y) - b(x, y) | (s + x), (t + y) \in D_f \, ; (x, y) \in D_b\} \quad (41)$$

$$(f \oplus b) = MAX\{f(s - x, t - y) + b(x, y) | (s - x), (t - y) \in D_f \, ; (x, y) \in D_b\} \quad (42)$$

where D_f and D_b are the domains of f and b. Several combinations of these basic operations are possible.

Werman et al.[12] introduced the use of operations of mathematical morphology for texture analysis. Using a straight line segment of length r and angle w as a structuring element for an operator O, a matrix $M_O(r, \omega)$ can be obtained. The elements of this matrix are computed by summing the gray-values of the output image, resulting from the use of operator O, with a structuring element defined by r, and ω, on the input image. Werman et al. composed matrices for different operations, e.g., dilation, and erosion. Several texture features can be derived from the matrices. The method, thus described, has one major drawback: it is not invariant under linear gray-scale transformations. This means that the texture features are dependent on the mean and variance of the gray-levels in the image. We developed a number of features which are by definition invariant under linear gray-scale transformations.

For this study we concentrated on two morphological gradients: the difference between the dilated and the eroded version of an image:

$$g = (f \oplus b) - (f \ominus b) \quad (43)$$

and the top-hat transformation,

$$h = f - \{(f \ominus b) \oplus b\}. \quad (44)$$

23

By varying the size, shape, and orientation of the structuring element, different aspects of the image texture are enhanced. The first gradient brings out the maximum variation of gray-levels within a given neighborhood thus highlighting sharp gray-level transitions. The top-hat transformation extracts white objects from the image, the shape, size, and orientation of which are dependent on the structuring element.

We used flat structuring elements i.e., $b(x,y)=0$. To quantify the information in the image after the application of the morphological gradients, the gray-level values of the pixels in the transformed image are summed. Since morphological gradients are intrinsically invariant under a uniform increase in image intensity, this property also holds for the derived texture features. In order for the derived texture features to be insensitive to a global multiplicative gray-scale change, the results of the morphological gradients with different structuring elements are combined into ratios. In this study, the following features were used

$$GR_i = \frac{\sum g_{\Box_{i,i}} - \sum g_{O_i}}{\sum g_{\Box_{i,i}} + \sum g_{O_i}} \tag{45}$$

$$GSS_{i,j} = \frac{\sum g_{\Box_{i,i}} - \sum g_{\Box_{j,j}}}{\sum g_{\Box_{i,i}} + \sum g_{\Box_{j,j}}} \tag{46}$$

$$GSR_{i,j} = \frac{\sum g_{O_i} - \sum g_{O_j}}{\sum g_{O_i} + \sum g_{O_j}} \tag{47}$$

$$GO_{i,j} = \frac{\sum g_{\Box_{i,j}} - \sum g_{\Box_{j,i}}}{\sum g_{\Box_{i,j}} + \sum g_{\Box_{j,i}}} \tag{48}$$

where O_i and $\Box_{i,j}$ are, respectively, a circular structuring element with radius i, and a rectangular structuring element of size $i \times j$. A comparable set of features, *HR, HSS, HSR,* and *HO,* can be defined for the top-hat transformation h. The features *HR* and *GR* are sensitive to the roundness of structures and edges, respectively. Whereas *HSS* and *HSR* assess the size of structures with respectively square and round structuring elements, *GSS* and *GSR* perform the same function for edges. The orientation of structures and edges present in the image are quantified by *HO* and *GO.*

Invariance under strict monotonic increasing gray-level transformation
The morphological gradient features are defined in such a way, as described above, that they are invariant under linear gray-level transformations. However, they are not invariant under strict monotonic increasing gray-level transformations.

References

1. Weszka JS, Dyer CR, Rosenfeld A. A comparative study of texture measures for terrain classification. IEEE Trans Syst Man Cybern 1976;6:269-285.
2. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern 1973;3(6):610-621.
3. Gonzalez RC, Woods RE. Digital Image Processing. (1st ed.) Reading, Massachusetts: Addison-Wesley Publishing Company, 1992.
4. Bajcsy R. Computer description of textured surfaces. Proc 3rd Int Joint Conf on Artificial Intelligence 1973;572-579.
5. Katsuragawa S, Doi K, MacMahon H. Image feature analysis and computer-aided diagnosis in digital radiography: classification of normal and abnormal lungs with interstitial disease in chest images. Med Phys 1989;16(1):38-44.
6. Liu S-S, Jernigan ME. Texture analysis and discrimination in additive noise. CVGIP 1990;49:52-67.
7. Pentland AP. Fractal-based description of natural scenes. IEEE Trans Pattern Anal Machine Intell 1984;6:661-674.
8. Peleg S, Naor J, Hartley R, Avnir D. Multiple resolution texture analysis and classification. IEEE Trans Pattern Anal Machine Intell 1984;6:518-523.
9. Mandelbrot BB. The fractal geometry of nature. New York: W.H. Freeman, 1977.
10. Lynch JA, Hawkes DJ, Buckland-Wright JC. Analysis of texture in macroradiographs of osteoarthritic knees using the fractal signature. Phys Med Biol 1991;36(6):709-722.
11. Serra J. Image analysis and mathematical morphology (vol 1). London: Academic Press, 1982.
12. Werman M, Peleg S. Multiresolution texture signatures using min-max operators. Int Conference on Pattern Recognition 1984;97-99.

Part I:

Simulation Studies

Chapter 3

Texture analysis in radiographs: the influence of MTF and noise on the discriminative ability of texture features

J.F. Veenland[1,2], J.L. Grashuis[1,2], E.S. Gelsema[1]

[1] Department of Medical Informatics,
[2] Department of Radiology, Erasmus University Rotterdam, The Netherlands

Abstract

Tissue structures, represented by textures in radiographs, can be quantified using texture analysis methods. Different texture analysis methods have been used to discriminate between different aspects of various diseases in primarily X-rays of chest, bone, and breasts. However, most of these methods have not specifically been developed for use on radiographs. Certain characteristics of the radiographic process, e.g., noise and blurring, influence the visible texture. In order for a texture analysis method to be able to discriminate between different underlying textures, it should not be too sensitive to such processes as image noise and blur.

In this study, we investigated the sensitivity of four different texture analysis methods to image noise and blur. First a baseline measurement was performed of the discriminative performance of the Spatial Gray-Level Dependence method, the Fourier Power Spectrum, the Fractal Dimension, and the Morphological Gradient Method on images, which were not affected by radiographic noise and blur. Two types of images were used: fractal images and Brodatz textures. Whereas the Brodatz images represent very different textures, the differences between the fractal images are more gradual. We assume that the behavior of the different texture analysis methods on the fractal images is representative for their performance on radiologic textures. On these two types of images we simulated the effect of four different noise levels and the effect of two different modulation transfer functions, corresponding with different screen-film combinations. The influence on the discriminative performance of the four texture analysis methods was evaluated.

The influence of noise on the discriminative performance is, as expected, dependent on the image type used: the discrimination of more gradually different images, such as the fractal images, is already lowered for relatively low noise levels. In contrast, when the images are more different, only high noise levels decrease the discriminative performance. The discriminative power of the Morphological Gradient Method is least affected by image blur.

We conclude that the discriminative performance of the Morphological Gradient Method is superior to that of other methods in circumstances which mimic the conditions prevailing in radiographs.

Keywords: *texture analysis, radiographs, modulation transfer function, noise*

1. Introduction

Already in 1972, the potential of texture analysis to discriminate between different aspects of diseases in radiographs was demonstrated[1]. Since then, various texture analysis methods have been applied to distinguish between different disease categories. For the classification of pulmonary diseases in chest radiographs, features derived from the Fourier power spectrum[1,2] and from the spatial gray-level dependence matrices[2,3] were employed. Katsuragawa et al.[4] defined two features using the Fourier power spectrum to detect interstitial disease in chest radiographs. The same features were used in spine radiographs to detect osteoporosis[5]. More recently, the fractal dimension has become a popular feature in quantifying textures in mammograms[6,7], and in bone radiographs[8-15]. Also, Sobel filters have been used in bone radiographs to generate features which correlate with the strength of bone[16].

However, most of the texture features which are used nowadays, such as the fractal dimension, features derived from the spatial gray-level dependence matrices, and from the Fourier power spectrum, have not been specifically developed for use on radiographs. Radiographs differ essentially from other imagery in the sense that they are dose limited: for patient safety, the exposure dose is kept as low as possible. This has two degrading effects on image quality. In the first place, the quantum mottle. Quantum mottle, the most important source of noise in radiographs, is the result of statistical fluctuations in the number of X-ray quanta per unit area absorbed by the intensifying screen[17,18]. Second, the intensifying screens which are used in front of the film in order to enhance the light output, have a blurring effect. The photons, generated in the intensifying screen by the absorption of X-ray quanta, are scattered in the screen before being absorbed by the film. Due to this scattering process, a point in the object will be depicted as a blob in the image. This blurring effect results in reduced spatial resolution, contrast, and sharpness. In the spatial frequency domain, this scattering effect is described by the Modulation Transfer Function (MTF)[18]. The MTF quantifies the relative transmittance of the spatial frequencies occurring in the depicted object.

These two effects, noise and blurring, are inherent to radiographic imaging. Therefore, in order to be able to quantify the characteristics of the underlying texture, the features used should be relatively insensitive to noise and blurring. Since both processes alter the visible texture, it is to be expected that all texture analysis methods are more or less affected. However, since different texture analysis methods are based on different image models, the sensitivity of the features derived by the different methods may vary.

Table 1. *List of all the texture analysis features used.*

	SGLDM	FM	FD	MGM
	Co-occurrence	Fourier	Fractal	Morphological Gradient
1	$ASM_{1,0^\circ}$	$P_{1,2}$	PSM_1	GR_5
2	$CON_{1,0^\circ}$	$P_{2,4}$	PSM_2	GR_7
3	$COR_{1,0^\circ}$	$P_{4,8}$	PSM_3	GR_9
4	$DMO1_{1,0^\circ}$	$P_{8,16}$	PSM_4	GR_{11}
5	$ENT_{1,0^\circ}$	$P_{0^\circ,45^\circ}$	BM_1	$GSS_{7,3}$
6	$ASM_{2,0^\circ}$	$P_{45^\circ,90^\circ}$	BM_2	$GSS_{9,5}$
7	$CON_{2,0^\circ}$	$P_{90^\circ,135^\circ}$	BM_3	$GSS_{11,7}$
8	$COR_{2,0^\circ}$	$P_{135^\circ,180^\circ}$	BM_4	$GSR_{7,3}$
9	$DMO1_{2,0^\circ}$	RMS		$GSR_{9,5}$
10	$ENT_{2,0^\circ}$	FMO		$GSR_{11,7}$
11	$ASM_{4,0^\circ}$	f_1		$GO_{7,1}$
12	$CON_{4,0^\circ}$	f_2		$GO_{9,3}$
13	$COR_{4,0^\circ}$	f_3		$GO_{11,3}$
14	$DMO1_{4,0^\circ}$	f_5		$GO_{11,5}$
15	$ENT_{4,0^\circ}$	f_6		$GO_{13,5}$
16	$ASM_{1,90^\circ}$	f_9		HR_5
17	$CON_{1,90^\circ}$	f_{10}		HR_7
18	$COR_{1,90^\circ}$	f_{11}		HR_9
19	$DMO1_{1,90^\circ}$	f_{12}		HR_{11}
20	$ENT_{1,90^\circ}$	f_{13}		$HSS_{7,3}$
21	$ASM_{2,90^\circ}$	f_7		$HSS_{9,5}$
22	$CON_{2,90^\circ}$	f_{17}		$HSS_{11,7}$
23	$COR_{2,90^\circ}$	f_{18}		$HSR_{7,3}$
24	$DMO1_{2,90^\circ}$	f_{19}		$HSR_{9,5}$
25	$ENT_{2,90^\circ}$	f_{20}		$HSR_{11,7}$
26	$ASM_{4,90^\circ}$	f_{21}		$HO_{7,1}$
27	$CON_{4,90^\circ}$	f_{25}		$HO_{9,3}$
28	$COR_{4,90^\circ}$	f_{26}		$HO_{11,3}$
29	$DMO1_{4,90^\circ}$	f_{27}		$HO_{11,5}$
30	$ENT_{4,90^\circ}$	f_{28}		$HO_{13,5}$

In a simulation study, using synthetic fractal images with known dimension, we found that the fractal dimension, as estimated by different methods, is seriously influenced by both processes[19]. Therefore, it seems likely that the discriminative performance of the fractal dimension, as estimated by different methods, will be hampered by both processes. This limiting effect on the discriminative performance will also be present for other texture analysis methods. By quantifying the effect of noise and blurring on the discriminative performance of the texture features and methods, an indication of the suitability of the different features for texture analysis studies in radiographic images is given.

In this paper we study the effect of certain radiographic imaging characteristics, namely noise and image blur, on the discriminative performance of four different texture analysis methods. First a baseline measurement is performed of the discriminative performance of, the features of, the Spatial Gray-Level Dependence method, the Fourier Power Spectrum, the Fractal Dimension, and the Morphological Gradient Method. Then the influence of noise and the modulation transfer function on the discriminative performance of the features and of each method is investigated using images on which these processes are simulated.

2. Texture analysis methods

In this study, we used four different texture analysis methods: the Spatial Gray-Level Dependence Method, the Fourier Method, two fractal dimension estimation methods (the Power Spectrum Method and the Blanket Method), and the Morphological Gradient Method. These methods are described in Chapter 2. In the following text we will refer to the features generated by the different methods as the co-occurrence features, the Fourier features, the fractal features and the morphological gradient features, respectively. An overview of all features used is given in Table 1.

Spatial Gray-Level Dependence Method
Firstly, the histogram of each image was equalized to four gray-levels. Then, for the computation of the intermediate matrix, three distances d were selected: values of 1, 2, and 4 pixels were used. For the angle θ values of 0° and 90° were taken. As the last step, the five features, as described in Chapter 2 were computed for each distance d and angle θ.

a)

b)
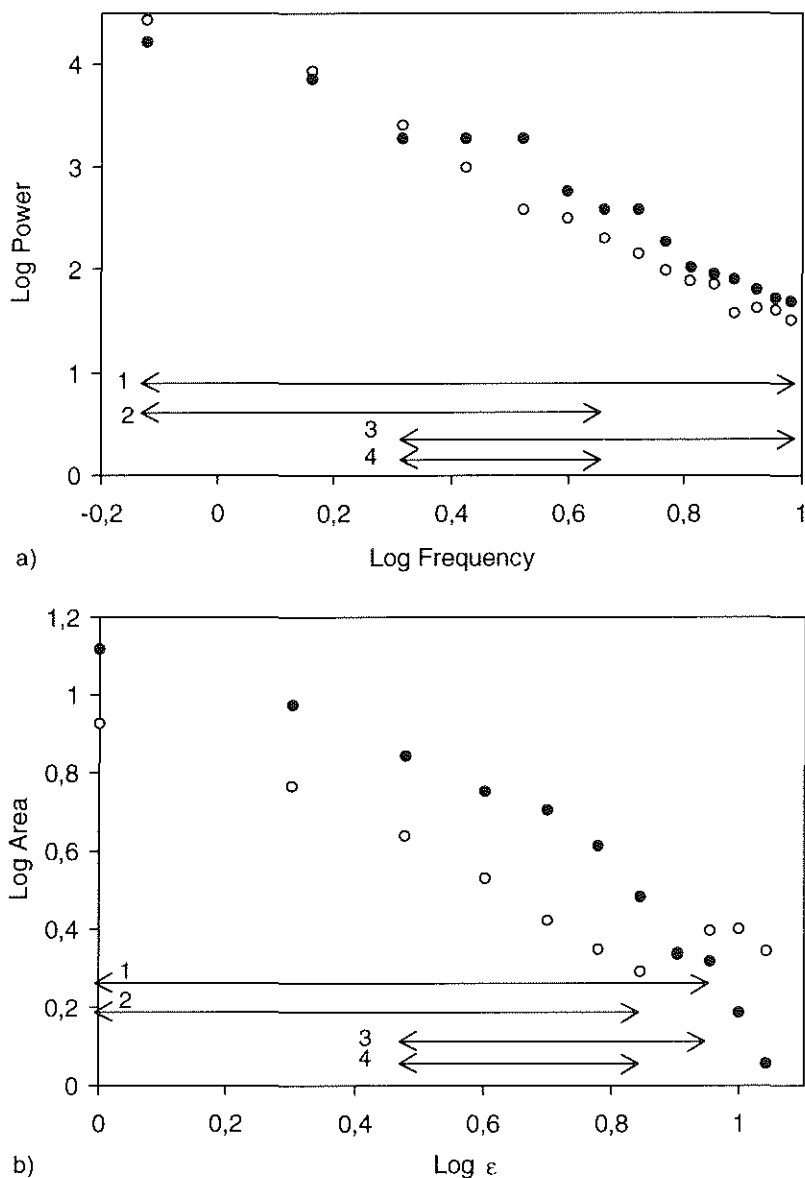
Fig 1. *Log-log plots as used for the fractal dimension estimation by the Power Spectrum Method (a) and the Blanket Method (b). The results for a subimage of the Brodatz (●) and fractal (o) image type are shown. The ranges we used in linear regression analysis to obtain four estimates are indicated at the bottom of the figures.*

Fourier Method

The three groups of features as described in Chapter 2 were used. For the first group of features four pairs of radii and angles were chosen, resulting in the following set of features: $P_{1,2}$, $P_{2,4}$, $P_{4,8}$, $P_{8,16}$, $P_{0°,4°}$, $P_{45°,90°}$, $P_{90°,135°}$, and $P_{135°,180°}$.
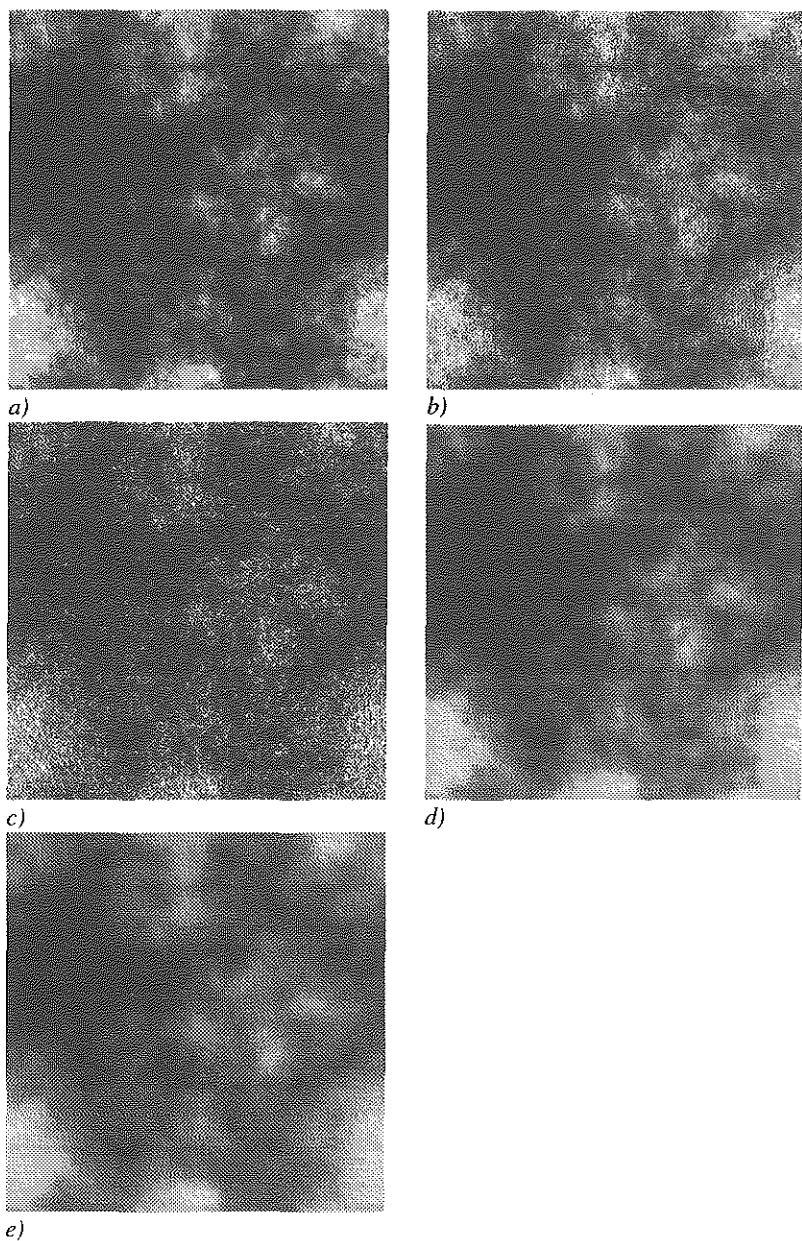
Fractal dimension

Two different fractal dimension estimation methods, as described in Chapter 2, were used: the Power Spectrum Method and the Blanket Method. Both methods are based on the same principle: an image characteristic is measured as a function of a scale parameter. The two methods differ in their definition of the image characteristic and the scale parameter. The relation between these two quantities can be described by a straight line in a log-log domain. The slope, computed by linear regression analysis, is linearly related to the fractal dimension. It is well recognized that the fractal behaviour of natural images is limited to a certain range of the scale parameter. Therefore, the linear regression analysis should not take all data-points, deduced from all available resolution scales, into account: an upper and lower boundary must be specified for the scale parameter. For the Power Spectrum Method and the Blanket Method, four different ranges were selected for regression analysis as indicated in Fig. 1(a) and 1(b), respectively, resulting in 8 fractal features, PSM_1, PSM_2, PSM_3, PSM_4, BM_1, BM_2, BM_3, and BM_4.

Morphological Gradient Method

This method was developed by ourselves especially for use in radiographs. The features are defined in Chapter 2, the different sizes of the structuring element used, are given in Table 1.

3. Methods

In order to study the influence of noise and image blur on the discriminative performance of the four texture analysis methods, simulation experiments were performed on textures which were not affected by radiographic noise and blur. On these images the effects of noise and MTF were simulated as follows. Each image is regarded to correspond with the X-ray intensity distribution in front of the screen.

Fig 2. The simulation of noise and MTF on the fractal image with dimension 2.5: original image (a), noisy images with SNR 10 (b) and 1 (c), blurred images with the MTF corresponding to Lanex Fine/0G (d), and Lanex Regular/0G (e).

36

The effect of quantum noise is then simulated using Gaussian noise with various signal-to-noise ratio levels; the effect of photon scattering in the screen-film system is simulated by using a Modulation Transfer Function. The resulting output image corresponds with the intensity distribution of the light in front of the film. The texture analysis methods are applied to the output images. In this way, effects of the characteristic film curve, which have been shown to influence texture features[20], are not taken into account.

Images

In order to perform a baseline measurement, textures which were not affected by radiographic noise and image blur, were needed. We used two types of images: fractal images and Brodatz textures[21]. The Brodatz textures have been widely used for the development and testing of texture analysis methods[22,23]. However, the Brodatz textures are rather dissimilar. In radiographs, the differences between textures are usually more gradual. Therefore we also used an image type, which is characterized by gradual differences between the different textures: fractal images. The fractal images were generated using the spectral synthesis method[24]. By this method, a power spectrum is generated which corresponds to an ideal fractal image with a given dimension. After the addition of Gaussian noise, an inverse Fourier transformation is performed which yields a fractal image. Five images (256x256x8 bits) were synthesized with fractal dimension 2.1, 2.3, 2.5, 2.7, and 2.9, respectively. Of the Brodatz textures, five images were selected: D17, D19, D28, D84, and D92[21].

All images were pre-processed in order to remove differences in first order statistics such as the mean gray-value and gray-level variance. An example of both image types is shown in Figs. 2(a) and 3(a): the fractal image with dimension 2.5 and the Brodatz image D28, respectively.

Noise

Usually, the signal-to-noise ratio (SNR) is defined[25] as the ratio between the average energy of the signal s and the average energy of the noise n:

$$SNR \equiv \frac{\sum\sum s^2}{\sum\sum n^2} = \frac{VAR(s) + AVG^2(s)}{VAR(n) + AVG^2(n)} \tag{1}$$

However, when images differ in average gray-level, the same SNR can only be reached by adjusting the average noise energy. Consequently, the contribution of noise to the texture is not comparable.
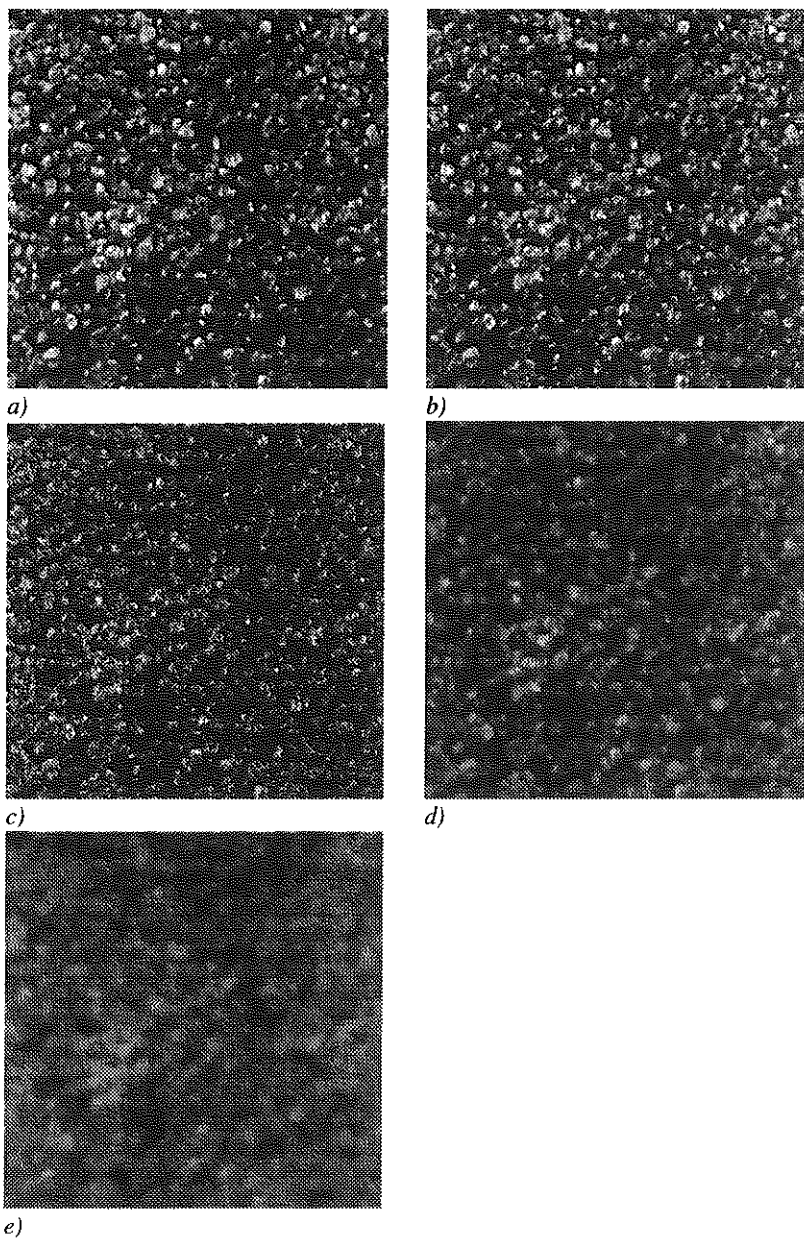
Fig 3. The simulation of noise and MTF on the Brodatz image D28: original image (a), noisy images with SNR 10 (b) and 1 (c), blurred images with the MTF corresponding to Lanex Fine/0G (d), and Lanex Regular/0G (e).

Therefore, we used the definition of the signal-to-noise ratio (SNR) as the ratio between the variances of the signal and the noise[26].

Using Gaussian zero mean noise, for each image four different noisy images were generated, with SNRs of 100, 10, 1, and 0.1. The effect of noise on the fractal and Brodatz images is shown in Figs. 2 and 3 for SNRs of 10 (b) and 1 (c), respectively.

Modulation Transfer Function

The Modulation Transfer Function (MTF) can be used to describe the spatial resolution characteristics of screen-film systems. Whereas the lower spatial frequencies of the object spectrum projected on the radiograph are only slightly attenuated, the transfer of the higher frequencies is seriously reduced. One of the causes is the scattering of light photons in the screen-film system. The influence of the scattering process can be spatially described by the point spread function (PSF). The modulus of the Fourier transform of the PSF is the modulation transfer function. We modelled the MTF using a logit function as described by Bencomo et al.[27]:

$$logit\ MTF(f\ ) = \log \{\ MTF(f\ )\ /\ (1 - MTF(f\ ))\ \} \qquad (2)$$

By logit analysis, the MTF of screen-film systems can be characterized by two parameters, *a* and *b*. These parameters describe the straight line which results when *logit MTF* is plotted against the log of the frequency *f*.

$$logit\ MTF(f\ ) = b \log (f\ ) + a \qquad (3)$$

A complete analysis of the influence of the MTF would require the exploration of the two-dimensional *(a,b)*-space. Since our main goal is to demonstrate the effect of the MTF, we used the parameters of two commonly used screen-film combinations, the Lanex Fine/0G and Lanex Regular/0G of Kodak, to model a relatively sharp and unsharp screen-film system, respectively. Figure 4 shows the MTF of the two screen-film combinations.

In our simulation experiments, we assumed that the absorption of one X-ray quantum results in the emission of a fixed number of light photons. The scattering process of these photons was simulated by convolving the input images with the point spread functions of the different screen-film systems. This convolution was implemented in the Fourier domain using the transfer functions corresponding to the two different screen-film combinations, assuming a sampling distance of 50 µm. In Figs. 2 and 3 the blurring effect of the two MTFs corresponding with Lanex
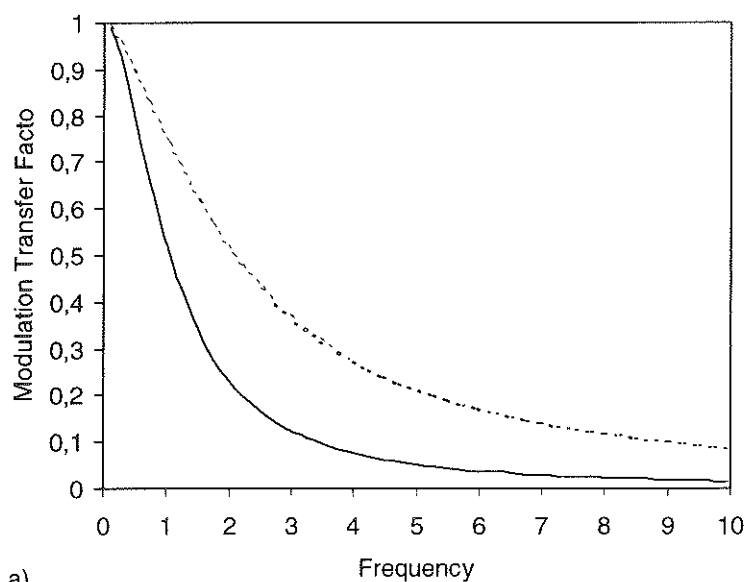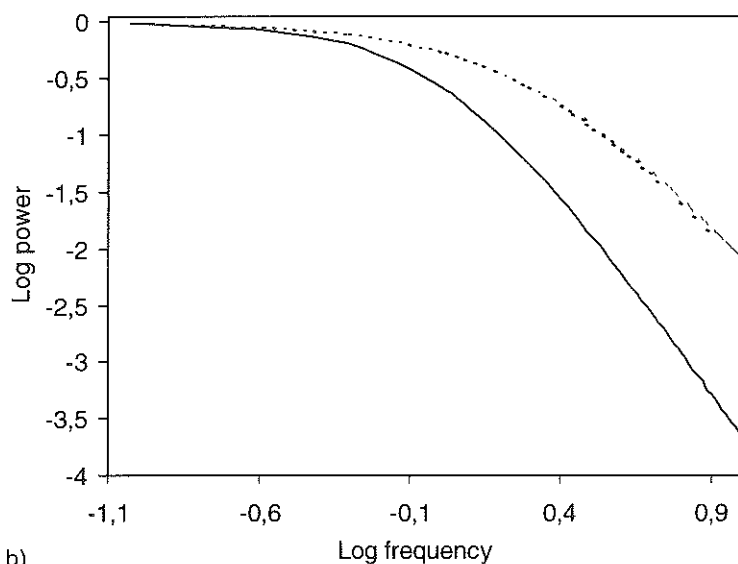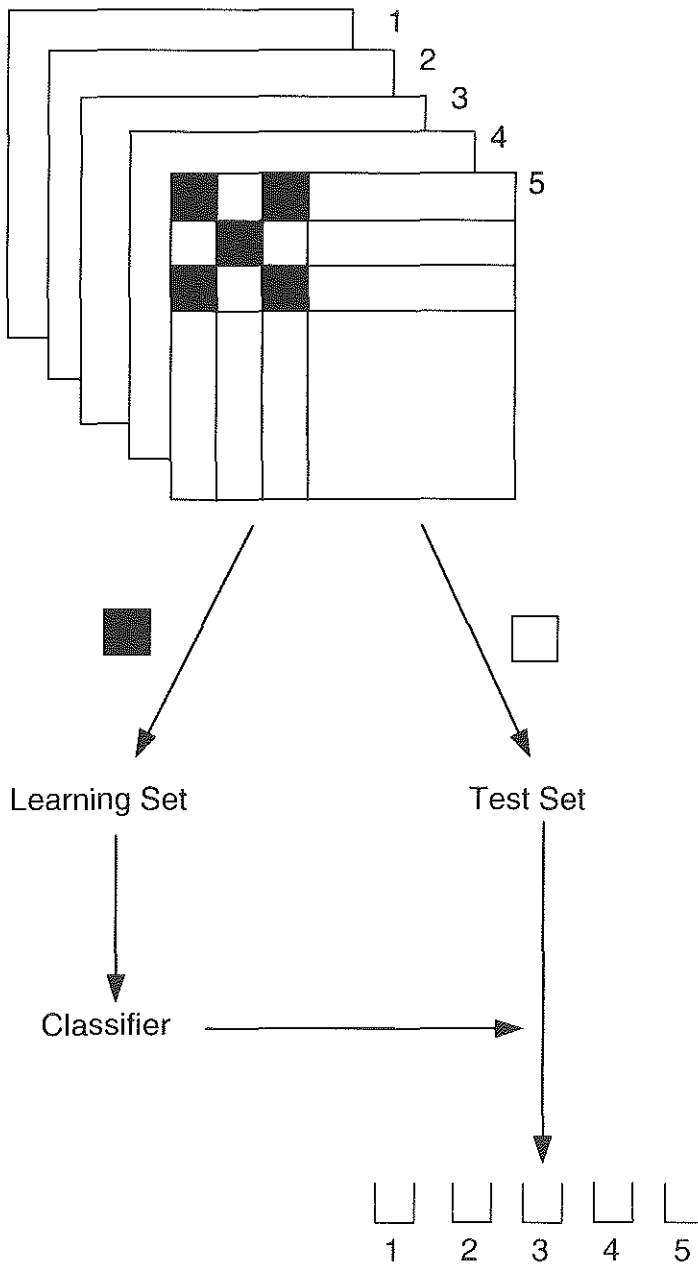
Fig 4. *(a) Modulation Transfer Functions computed by the logit model and (b) the relation between the log Fourier Power and log frequency of the model of the two screen-film combinations used: Lanex Regular/0G (solid line) and Lanex Fine/0G (dashed line).*

Fig 5. *Schematic representation of the classification experiment for one image type.*

Fine/0G (d) and Lanex Regular/0G (e) of Kodak, is shown for the image with fractal dimension 2.5 and the Brodatz image D28, respectively.

Classification

For the classification experiments two types of images, fractal and Brodatz, were used. In Fig. 5 the design of the classification experiment is schematically represented for one image type. Each image type contains five images, each image constitutes one class. Therefore, per image type five classes can be discerned. Each image was subdivided into 64 non-overlapping subimages of 32x32 pixels. The objective of the decision function is to classify the subimages according to their class, which is the original image from which they were extracted. For each of the five images, c.q. classes, of one image type the set of subimages was divided into two subsets (shown as black and white in Fig. 5). By merging the subsets, one for each class, the learning set and the test set for the decision function were created. The texture features, described in Section 2, were computed for each subimage. Subsequently, for each image type and each texture analysis method, the best combination of four features was selected using the learning set. This selection was based on the F-ratio, which is defined as the ratio of the between-classes variance and the within-classes variance. The best combination of four texture features was selected in a stepwise procedure: after the selection of the feature with the highest F-value, a regression of the set of candidate features on the set of selected features is performed, thus taking correlations between the features into account.

The four features thus selected were included in the decision function: the minimum Mahalanobis distance classifier[28]. This classifier assumes a normal distribution of the features. We used the linear variant for which, in addition, equality of covariance matrices for the different classes is assumed, since the number of objects per class is too limited, in view of the dimensionality of the feature space to arrive at a reasonably reliable estimate of the covariance matrix per class. The Mahalanobis distance from a pattern vector x to the mean vector m_i of class i is defined by

$$D_i = (x - m_i)^t C^{-1}(x - m_i) \tag{4}$$

In the learning phase, the mean vectors m_i are estimated from the corresponding sub-images in the learning set; the single covariance matrix C is estimated by pooling the objects from the different classes. The performance of the classifier is assessed by determining for each subimage of the test set the class i for which D_i is minimum and by evaluating the proportion of subimages of the test set which are

correctly classified. This proportion is called the Correct Classification Rate (CCR).

Using the ingredients listed above, four series of experiments were performed.

A. *Comparison of the different features.*
   The F-values of the different features within each method were compared. The features are listed in Table 1.

B. *Comparison of the different methods.*
   The discriminative performance of the four texture analysis methods on the two types of images was compared. The correct classification rate on the original, unperturbed images serves as a reference value for the other experiments.

C. *Influence of noise.*
   The effect of noise was simulated on the input images. Different signal-to-noise ratios were used: 100, 10, 1, and 0.1, respectively. For both image types, the classification experiments were performed on the resulting images.

D. *Influence of the MTF.*
   The effect of the Modulation Transfer Function was simulated on the input images. Two different MTFs were used, corresponding to two screen-film systems, Lanex Fine/0G and Lanex Regular/0G of Kodak, a relatively sharp and an unsharp screen-film system, respectively. For both image types, the classification experiments were performed on the resulting images.

4. Results

A. Comparison of the different features

In Fig. 6, the F-values of the different features of the Spatial Gray-Level Dependence Method are depicted. For the fractal images (Fig. 6(a)), the decrease of the F-value with increasing distance is prominent. The F-values for the two orientations are comparable, as are the F-values for the five different features (ASM, CON, COR, DMO1, and ENT). In the Brodatz images (Fig. 6(b)), a difference in orientation can be noted: the features measured under an angle $0°$ have higher F-values than those measured under $90°$. Although less distinct than in the fractal images, the F-values for the features measured at the smaller distance, $d=1$, are better discriminative than those for $d=2$ and $d=4$. Furthermore, the F-values for Angular Second Moment and Entropy are slightly higher than for the other features.

a)

b)
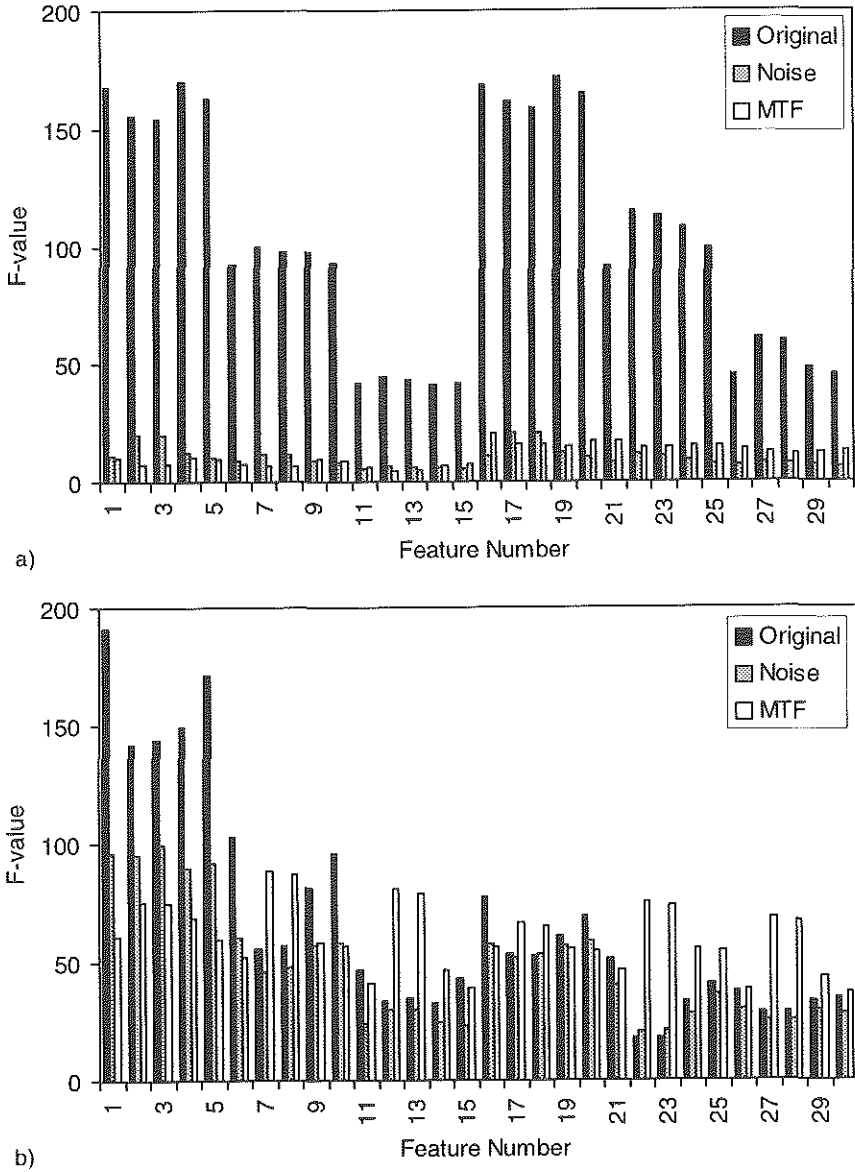
Fig 6. The F-values of the co-occurrence features in the fractal (a) and Brodatz(b) images: for the original images, the images with SNR 10, and the images after the simulation of the effect of the Modulation Transfer Function of the Lanex Regular/OG screen-film system.
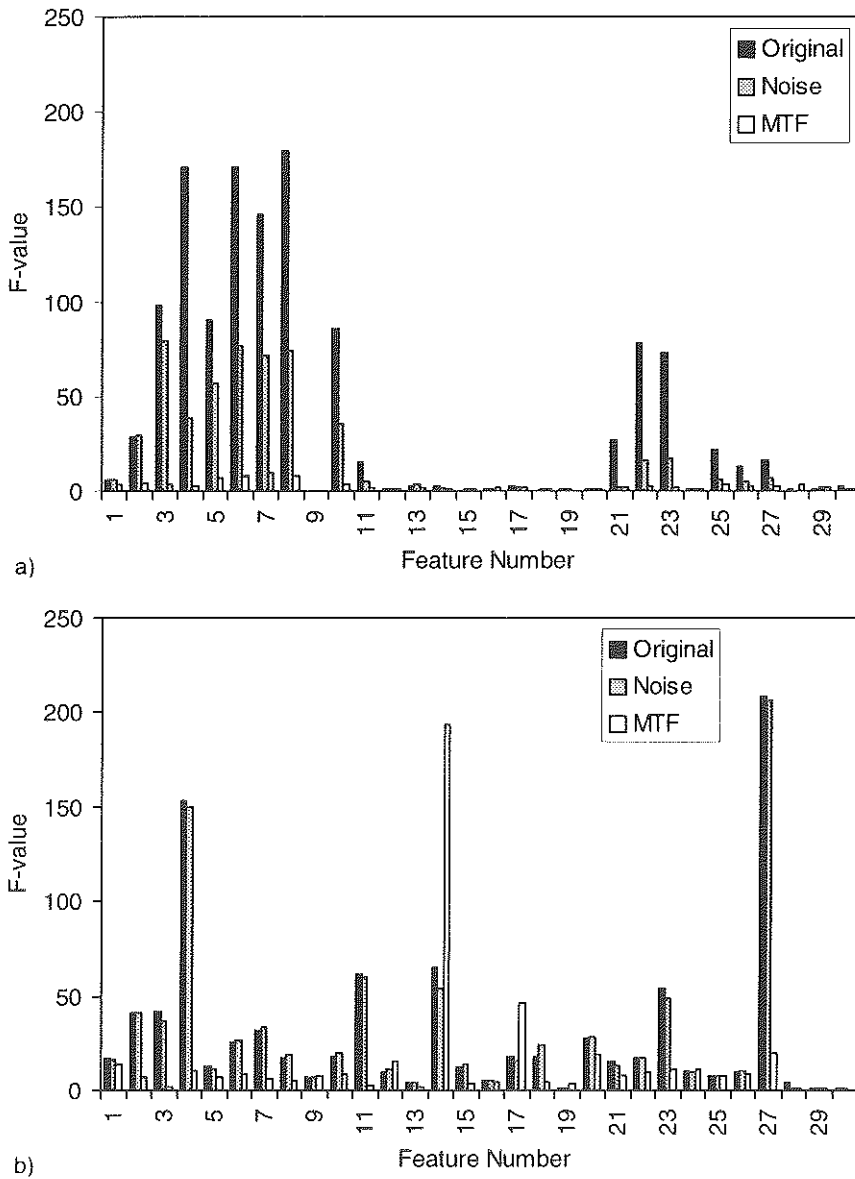
Fig 7. *The F-values of the fourier features in the fractal (a) and Brodatz(b) images: for the original images, the images with SNR 10, and the images after the simulation of the effect of the Modulation Transfer Function of the Lanex Regular/0G screen-film system.*

45
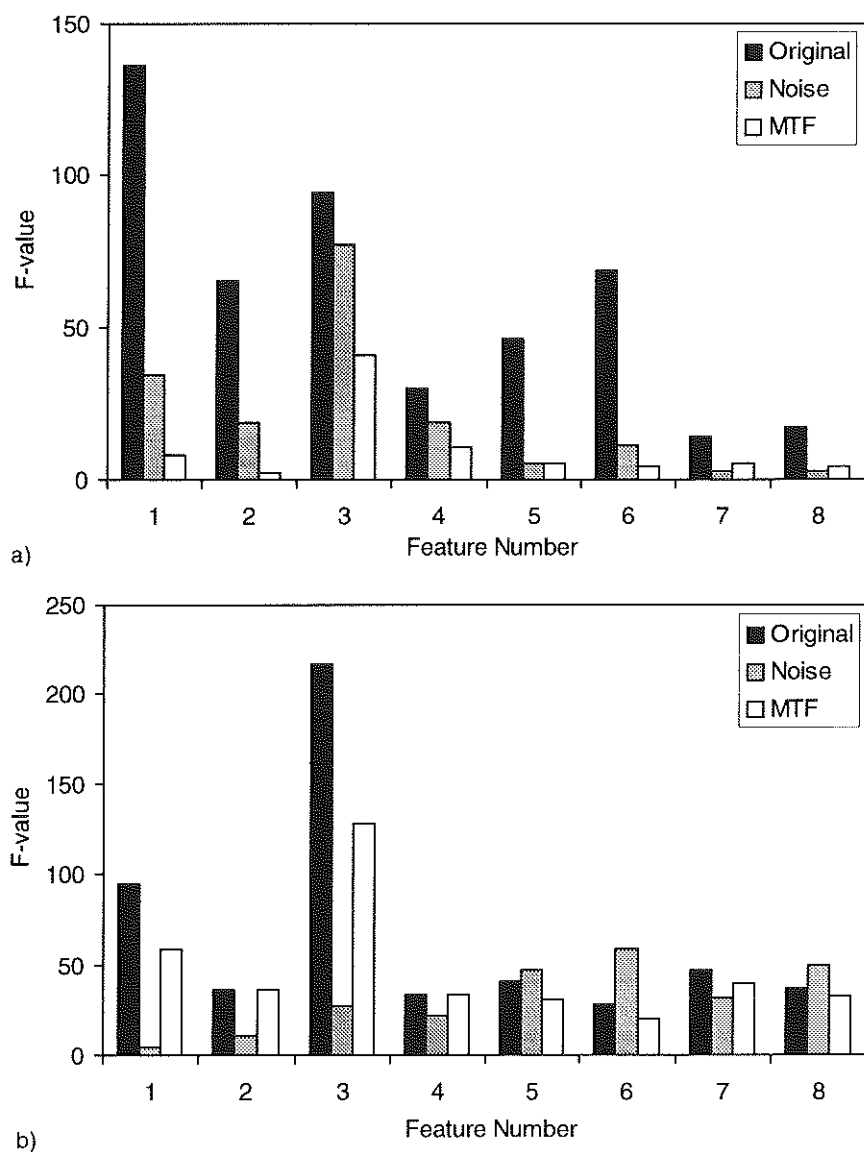
Fig 8. The F-values of the fractal features in the fractal (a) and Brodatz(b) images: for the original images, the images with SNR 10, and the images after the simulation of the effect of the Modulation Transfer Function of the Lanex Regular/0G screen-film system.

Fig 9. *The F-values of the morphological gradient features in the fractal (a) and Brodatz(b) images: for the original images, the images with SNR 10, and the images after the simulation of the effect of the Modulation Transfer Function of the Lanex Regular/0G screen-film system.*
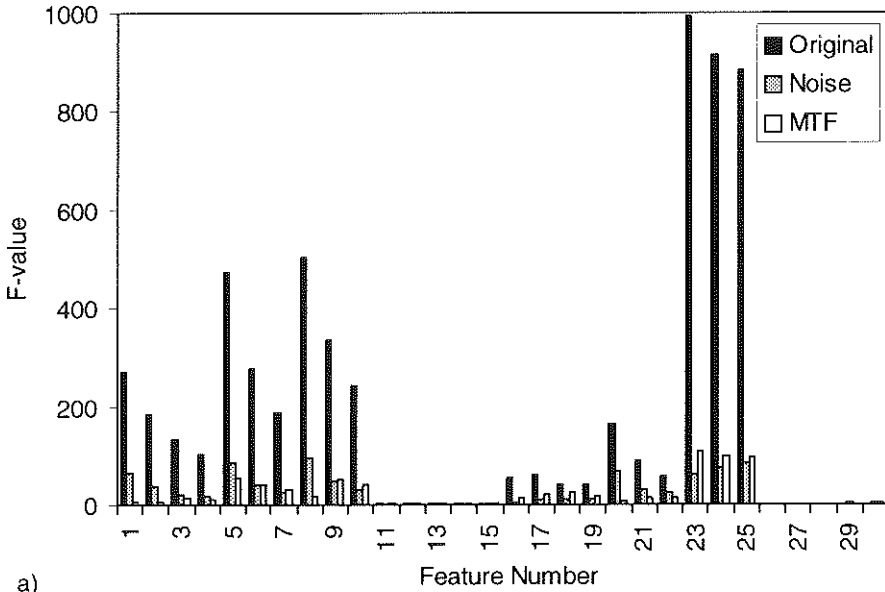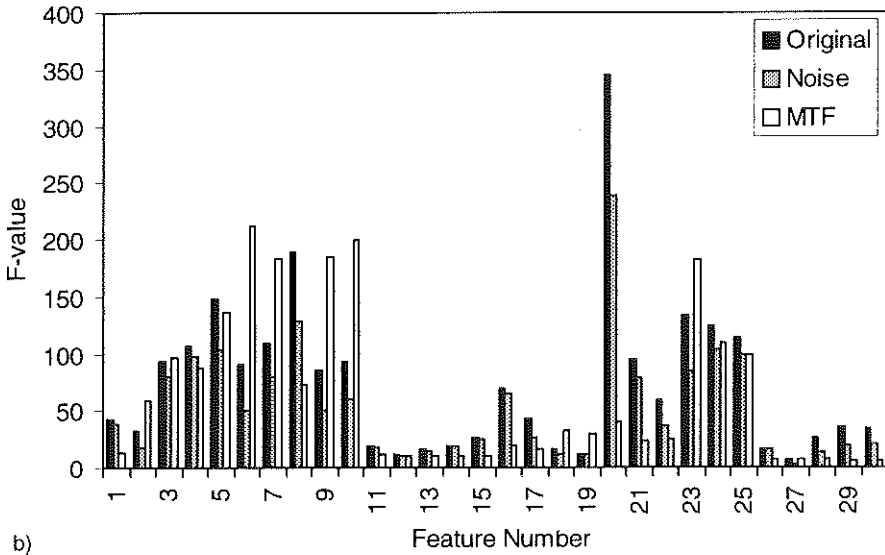
The F-values of the features generated by the Fourier method are shown in Fig. 7 for the fractal (a) and the Brodatz (b) images. For the fractal images, the first group of features, which quantifies the radial and angular power distribution, performs the best. The F-values of the local spectrum features of the third group (f_{11} up to f_{20}) are all very low, whereas some of the global spectrum features have moderate F-values. In contrast, in the Brodatz images, no distinction between the groups can be made.

For the two fractal dimension estimation methods, the F-values of the features are displayed in Fig. 8 for the fractal (a) and the Brodatz (b) images. In both image types, the Power Spectrum Method generates the features with the highest F-values.

In Fig. 9, the F-values for the features generated by the Morphological Gradient Method are shown for both image types: the fractal (a) and Brodatz (b) images. In both image types, the orientation features (11–15 and 26–30) have low F-values, and also the HR-features (16–19) perform poorly. In the fractal images, the performance of the other features is generally lower when larger structuring elements are used.

B. Comparison of the different methods
The discriminative performance of the four texture analysis methods on the two types of images was compared. In Fig. 10, the correct classification rates based on the best combination of four features, for each of the four methods, for the fractal and Brodatz images are shown. The classifier based on morphological gradient features achieves the highest score in both image types. In the Brodatz images, the classifier based on the features generated by the Fourier method has the lowest score, whereas in the fractal images the performance of the classifier based on the Fourier features, the co-occurrence features, and the fractal features is comparable. It should be noted that in the fractal images, the fractal features do not attain a higher score than those generated by the other methods. On the other hand, the fractal features perform rather well in the Brodatz images, although these are not specifically fractal in nature.

C. Influence of noise
The effect of noise was simulated on the input images. Different signal-to-noise ratios were used: 100, 10, 1, and 0.1, respectively. For both image types and for each method, the F-values of the individual features were determined on the resulting images. The effect of noise, exemplified by the signal-to-noise ratio of 10, on the F-value of the individual features of each method is shown in Figs. 6, 7, 8, and 9. It can be seen that on the fractal images, noise seriously decreases the F-values.
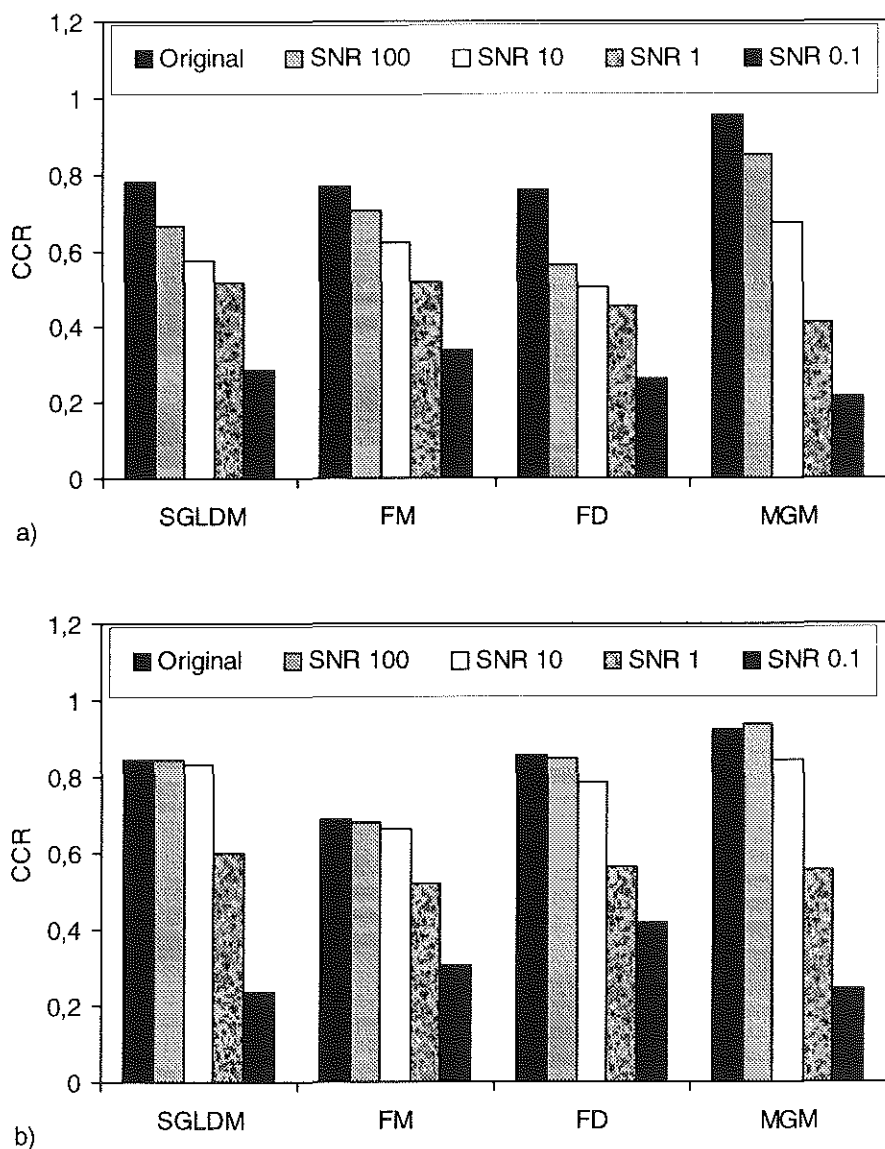
48
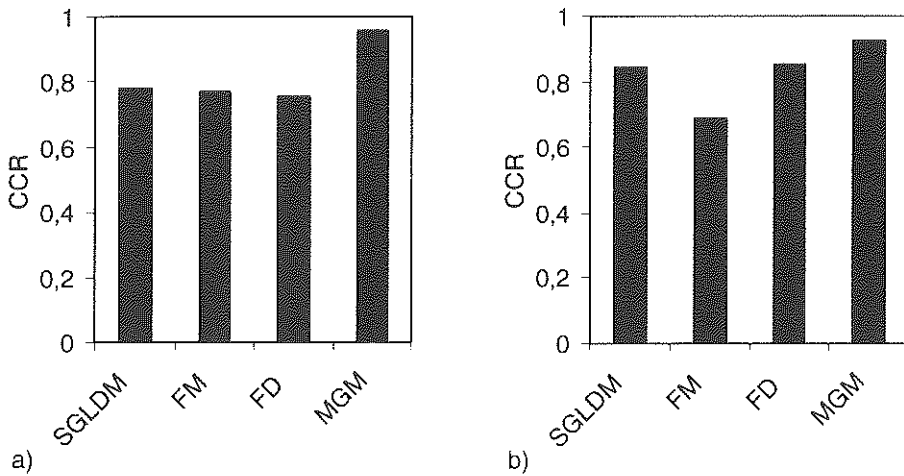
Fig 10. The correct classification rate (CCR) of four texture analysis methods on Fractal (a) and Brodatz (b) images (SGLDM = Spatial Gray-Level Dependence Method, FM = Fourier Method, FD = Fractal Dimension, MGM = Morphological Gradient Method).

The co-occurrence features and the morphological gradient features are the most sensitive to noise. On the other hand, when measured on the Brodatz images, most features show a less marked decrease in the F-values, especially the features of the Fourier method are rather insensitive to noise.

In the fractal images, the co-occurrence features still have the highest F-values for distance 1; there is no noticeable difference between the two different orientations. The features Contrast and Correlation perform slightly better than the other features (Fig. 6(a)). For the Brodatz images, the effect of noise is less marked. The features with direction 0° and distance 1 are most affected by noise, but still have the highest F-values (Fig. 6(b)).

The F-values of the features of the Fourier method are, especially in the Brodatz images (Fig. 7(b)) not much affected by noise. The fractal features as estimated by the Power Spectrum Method show in both types of images (Fig. 8) the highest F-values in a noisy environment. In Fig. 9 the effect of noise on the F-value of features generated by the morphological gradient method is shown.

The influence of noise on the correct classification rate for fractal (a) and Brodatz (b) images is depicted in Fig. 11. In the fractal images, a signal-to-noise ratio of 100 results already in a decrease in performance. On the contrary, in the Brodatz images the classification is not much hampered as long as the signal-to-noise ratio is 10 or higher, irrespective of the texture analysis method used.

a)

b)

Fig 11. *The influence of noise on the correct classification rate (CCR of the four texture analysis methods) in the fractal (a) and Brodatz (b) images (SGLDM = Spatial Gray-Level Dependence Method, FM = Fourier Method, FD = Fractal Dimension, MGM = Morphological Gradient Method).*
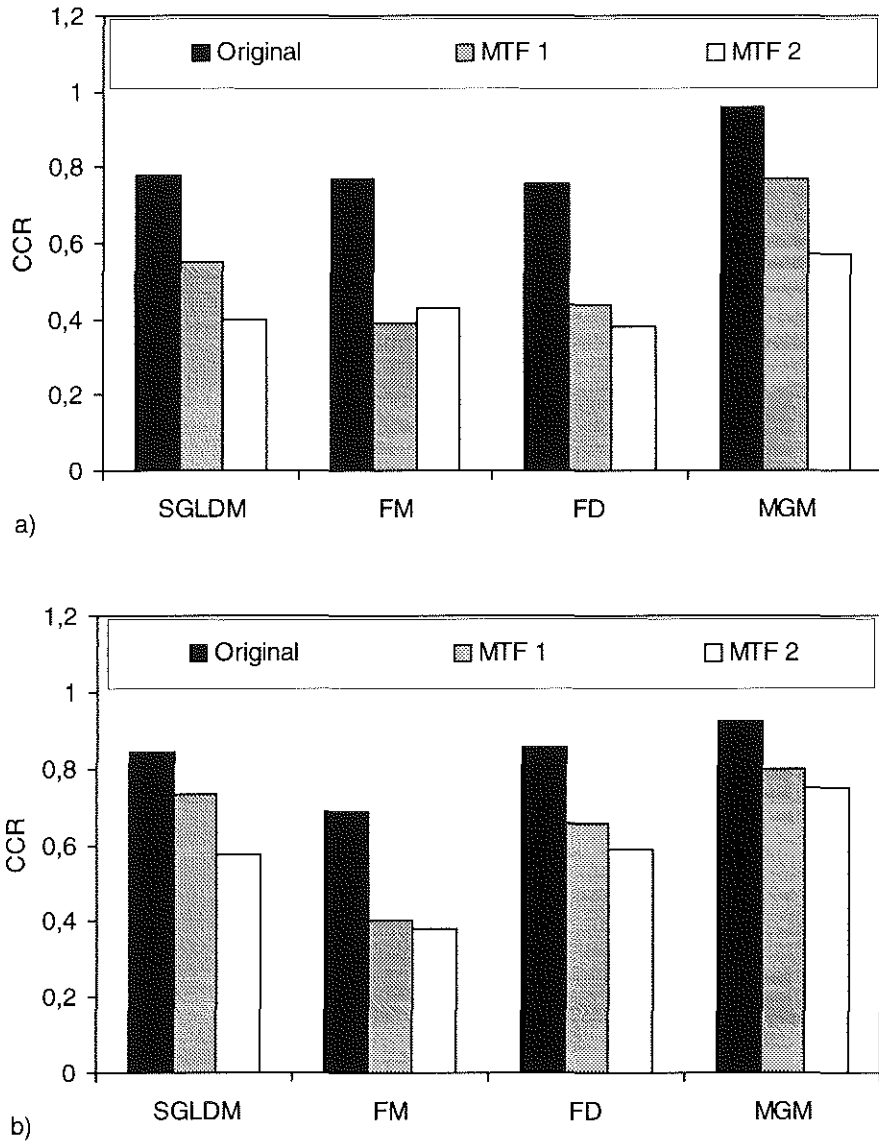
Fig 12. *The influence of the MTF on the correct classification rate (CCR of the four texture analysis methods) in the fractal (a) and Brodatz (b) images (SGLDM = Spatial Gray-Level Dependence Method, FM = Fourier Method, FD = Fractal Dimension, MGM = Morphological Gradient Method).*

51

This result is in accordance with the effect of noise on the individual features as described above. As shown above, for the fractal images the effect of a signal-to-noise ratio of 10, is rather detrimental for the discriminative performance of the features, irrespective of the method used. In contrast, in the Brodatz images, the performance of the features is not much hampered. In both image types, the MG-Method shows the highest discriminative performance when the signal-to-noise ratio is 10 or higher. In fractal images with a signal-to-noise ratio of 100, especially the classification based on the fractal features is hampered. For the other methods, the decrease in performance with lower signal-to-noise ratios is rather gradual, except for the ratio of 0.1. At that signal-to-noise level, the performance of the classifier is not much better than the result to be expected of a random classification.

D. Influence of the MTF

The influence of the MTF, corresponding with Lanex Regular/0G, on the F-values of the different texture features is shown in Fig. 6, 7, 8, and 9. In the fractal images, the F-values of the features of all methods are seriously decreased. In the Brodatz images, the decrease in the F-values is less marked; some features, especially the co-occurrence features and the morphological gradient features, show even higher values.

In the Brodatz images, the F-values of the co-occurrence features with distance 1 are reduced, whereas the F-values of features with distance 4 are enhanced. The resulting effect is that the variations present between the different orientations and distances in the original images, is decreased by the MTF as shown in Fig. 6(b). The features Contrast and Correlation seem to perform best. In Fig. 7 it can be seen that the F-values of the Fourier features are seriously reduced by the influence of the MTF, with the only exception, in the Brodatz images, of the features which determine the location of the spectrum peak. Figure 8(b) shows that three of the features determined by the Blanket Method, show, in contrast to the features generated by the Power Spectrum Method, a higher F-value, when tested on the Brodatz images. The performance of some of the morphological gradient features is increased by the influence of the MTF (Fig. 9(b)).

In Fig. 12, the influence of the MTFs corresponding with two screen-film combinations with a high and medium resolution (Lanex Fine/0G and Lanex Regular/0G of Kodak) on the discriminative performance is shown for fractal (a) and Brodatz (b) images. For all texture analysis methods, the performance of the classifier using the best subset of four features is decreased, as compared to the performance on the original images. For all methods, the performance is more

decreased in the fractal than in the Brodatz images. It is clear that the Fourier Method is, in both image types, the method most affected by image blur. It should be noted that the extent of the blur, whether it is caused by a high or medium resolution screen-film combination, seems irrelevant. This result is in accordance with the serious reduction in F-values due to the influence of the MTF, as described above for the individual Fourier features. The blur caused by the MTF corresponding with the high resolution screen-film system has a relatively small effect on the performance of the classifiers based on the co-occurrence and the morphological features. Especially in the Brodatz images, the Spatial Gray-Level Dependence Method still performs rather well despite the blur. In both image types, the Morphological Gradient Method is the least affected by image blur, irrespective of the resolution of the screen-film system modelled.

5. Discussion

In literature several texture analysis methods have been used to discriminate between different radiographic aspects of various diseases. However, the radiographic process has certain characteristics which influence the image texture. Therefore, in order for a texture analysis method to be able to discriminate between different textures in radiographs, the method should be relatively insensitive to these characteristics. In this paper, we focused on the influence of two aspects of the radiographic imaging chain, namely noise and the image blur caused by the intensifying screen. By simulating both processes on two types of images, using different levels of Gaussian noise and two different modulation transfer functions, the influence of noise and image blur on different texture analysis methods could be evaluated.

Comparison of the different methods
Several authors have compared the performance of texture analysis methods on different types of images. Weszka et al.[29] compared the performance of, among other methods, the Spatial Gray-Level Dependence Method, and the Fourier Method on aerial photographic terrain samples. They found that the co-occurrence features perform slightly better than the Fourier features. Conners et al.[30] used artificially generated textures to evaluate the performance of four texture analysis methods. The conclusion of their study was that the SGLDM is the most powerful texture analysis method of the four methods evaluated. Ohanian et al.[23] compared, on four types of images, four texture analysis methods, among which the Spatial

Gray-Level Dependence Method and fractal dimension estimation methods. In their experiments, fractal features performed best on fractal images, whereas on the other image types the co-occurrence features were more powerful. Liu et al.[22] evaluated the performance of Fourier features in additive noise on Brodatz images. They found correct classification rates as high as 97.5% in the original images. Interpreting this very high rate, it should be taken into account that the spectra of the images used were very different. The finding that additive noise had little influence on the discriminative performance must be interpreted also in that light.

From the results of these studies, it becomes clear that the performance of a texture analysis method is dependent on the type of image used. In this study, we used two types of images: fractal and Brodatz images. Whereas the Brodatz textures can be considered to be rather distinct, the differences between the fractal images are more gradual. Since the fractal dimension is a continuous variable, the fractal images are part of a continuum. Since the expression of a disease in radiologic images is usually a continuum, we assume that the results for fractal images are more relevant to radiographic textures.

Although the results of the different methods on the original, unperturbed images were primarily intended for use as a reference value for the influence of noise and the MTF, some comments can be made. In the first place, in both image types, the performance of the Morphological Gradient Method is superior to the other methods. Whereas the performances of the other methods for the fractal image type are comparable, in the Brodatz images the Spatial Gray-Level Dependence method and the fractal method perform better than the Fourier method. We cannot reproduce the very high correct classification rate as mentioned by Liu et al.[22]. However, we used different images, and a different classifier.

Influence of noise
The discriminative performance of all texture features studied is impeded by noise. However, this effect is more prominent in the fractal image type, where the similarity between the original images is closer. The smaller texture differences that exists between the original images are lost in the additive noise. Of the four methods studied, the performance of the co-occurrence features is most affected by the addition of noise. This is to be expected since the co-occurrence features measure gray-level differences between pixels at very small distances apart. Noise will have a more profound effect on this type of features than on features which are based on larger neighbourhoods or even on the whole image as is the case in the other methods. In a previous study[19], we already showed that the performance of fractal dimension estimation methods is limited by noise. The range of dimensions,

as estimated by the different methods, is reduced, and therefore the ability to discriminate between images with different fractal dimensions or different textures is reduced as well.

When the performance of the methods as a whole is evaluated, the same dependence of the effect of noise on the image type can be observed. In the Brodatz images, the performance of all methods is hardly influenced by noise as long as the signal-to-noise ratio is 10 or larger. In the case of lower signal-to-noise ratios, the performance of all methods decreases considerably. On the contrary, in the fractal images, even a very small amount of noise already has a degrading effect on the discriminative performance of all methods. This again can be explained by the larger dissimilarities between the Brodatz images compared to the fractal images. For both image types applies, that as long as the signal-to-noise ratio is above 10, the performance of the Morphological Gradient Method is superior to that of the other methods.

The claim by Liu et al.[22] that frequency domain features are less sensitive to noise, is not supported by our results. In this study, all methods are negatively influenced by noise. There is not one method which in both types of images for different signal-to-noise ratios performs consistently better than the other methods.

Influence of MTF

The effect of photon scattering in the intensifying screen, simulated by using the Modulation Transfer Function is more pronounced for the fractal image type. The explanation given for the effect of noise also holds for image blur: the textures of the fractal images are more similar. The textural differences that existed between the original images become smaller, or even disappear, by blurring the image.

For the co-occurrence features, the negative correlation between discriminative performance and distance is blurred out: the features have similar F-values for all distances. This can be explained by the blurring effect of the MTF, which increases the correlation between the gray-values of neighbouring pixels. Since the MTF has a blurring effect in all directions, the directional difference of the features is also mitigated. The F-values of the Fourier features are decreased by image blur, except for some of the local spectrum features in the Brodatz images. This effect depends on the choice of the Brodatz images. For example, the F-value of the fourteenth feature (f_5), the squared major peak frequency, is strongly enhanced by image blur: when the major peak frequencies of the original images are similar, blurring will further deteriorate the discriminative performance. However, when they are rather distinct, blurring will reduce the noise and therefore make a more exact localization of the peak frequency possible, resulting in an enhanced discriminative performance. We showed, in a previous study[19], that the discriminative power of

the fractal dimension as estimated by the Power Spectrum Method is independent of the MTF. However, this only holds when the image size is rather large. In the present study, using small subimages, the estimated dimension will vary considerably around the real dimension. The range of fractal dimensions as estimated by the Blanket Method is reduced by the influence of the MTF[19]. Therefore, the F-values of the Blanket Method features show a decrease due to the MTF in the fractal images. However, in the Brodatz images the F-values are increased under the influence of the MTF. This can probably be ascribed to the larger textural differences between the Brodatz images.

Comparing the performance of the methods as a whole, it is clear that in both image types the performance of the Morphological Gradient Method is the least affected by the MTF. The discriminative power of the Fourier Method is the most deteriorated due to the MTF in both image types.

In this study, we assume that radiographic images with the same characteristics, e.g., noise and MTF, are compared. Since all texture features we studied are influenced by these characteristics, features which are obtained from radiographs which differ in noise content or MTF cannot readily be compared. This is in agreement with a previous study[19], where we have shown that the fractal dimensions estimated in radiographs with different MTFs or noise properties cannot be compared. Our simulation model was confined to the effect of noise and the MTF. We are well aware of the fact that we left out several factors which may also influence texture measurements. Our aim was to illustrate the limiting effect, resulting from the MTF and noise, on the discriminative power of different texture analysis methods. Since these two processes already have a rather detrimental influence, other factors will only further deteriorate the discriminative power.

In this study, we did not try to find the best classifier, we did not aspire to give a classification recipe. The combination of features which has the best classification result is always dependent on the classes chosen. We did try to illustrate, the sensitivity of the different features and methods for the influence of noise and the modulation transfer function. Therefore, we hope that the reader who is planning to perform a study on radiographic textures, will use our results in deciding on which method or features could be useful given the radiographic characteristics of his (or her) material.

6. Conclusions

Summarizing, the following can be stated. In the original images the performance of the Morphological Gradient Method is superior to that of the other methods, especially if images from a continuum, like the fractal images, are to be classified.

All texture analysis methods are negatively influenced by noise, although the influence of noise depends on the image type. When the images are rather different, the discriminative performance of all texture analysis methods tested are robust for small quantities of noise. In contrast, when the images are more similar, the influence of additive noise on the discriminative power, even in small quantities, is detrimental. The discriminative performance of the Morphological Gradient Method is the least affected by the MTF. Since we assume that radiographic images are similar to fractal images because of the gradual differences between images, it can be expected that in radiographs the Morphological Gradient Method will show a discriminative performance superior to other texture analysis methods.

References

1. Sutton RN, Hall EL. Texture measures for automatic classification of pulmonary disease. IEEE Trans Comp 1972;C-21(7):667-676.
2. Kruger RP, Thompson WB, Turner AF. Computer diagnosis of pneumoconiosis. IEEE Trans Syst Man Cybern 1974;4:40-49.
3. Tully RJ, Conners RW, Harlow CA, Lodwick GS. Towards computer analysis of pulmonary infiltration. Invest Radiol 1978;13:298-305.
4. Katsuragawa S, Doi K, MacMahon H. Image feature analysis and computer-aided diagnosis in digital radiography: classification of normal and abnormal lungs with interstitial disease in chest images. Med Phys 1989;16(1):38-44.
5. Caligiuri P, Giger ML, Favus MJ, Jia H, Doi K, Dixon LB. Computerized radiographic analysis of osteoporosis: preliminary evaluation. Radiology 1993;186:471-474.
6. Caldwell CB, Stapleton SJ, Holdsworth DW, et al. Characterisation of mammographic parenchymal pattern by fractal dimension. Phys Med Biol 1990;35:235-247.
7. Priebe CE, Solka JL, Lorey RA, et al. The application of fractal analysis to mammographic tissue classification. Cancer Letters 1994;77:183-189.
8. Lynch JA, Hawkes DJ, Buckland-Wright JC. Analysis of texture in macroradiographs of osteoarthritic knees using the fractal signature. Phys Med Biol 1991;36(6):709-722.
9. Ruttimann ES, Webber RL, Hazelrig JB. Fractal dimension from radiographs of peridental alveolar bone. Oral Surg Oral Med Oral Path 1992;74:98-110.
10. Webber RL, Underhill TE, Horton RA, Dixon RL, Pope TL. Predicting osseous changes in ankle fractures. IEEE Eng Med Biol 1993;12:103-110.
11. Samarabandu J, Acharya R, Hausmann E, Allen K. Analysis of bone X-rays using morphological fractals. IEEE Trans Med Imaging 1993;12(3):466-470.

12. Caligiuri P, Giger ML, Favus MJ. Multifractal radiographic analysis of osteoporosis. Med Phys 1994;21(4):503-508.
13. Benhamou CL, Lespessailles E, Jacquet G, et al. Fractal organization of trabecular bone images on calcaneus radiographs. J Bone Miner Res 1994;9(12):1909-1918.
14. Buckland-Wright JC, Lynch JA, Rymer J, Fogelman I. Fractal signature analysis of macroradiographs measures trabecular organization in lumbar vertebrae of postmenopausal women. Calcif Tissue Int 1994;54:106-112.
15. Chen J, Zheng B, Chang Y-H, Shaw CC, Towers JD, Gur D. Fractal analysis of trabecular patterns in projection radiographs. Invest Radiol 1994;29(6):624-629.
16. Caldwell CB, Willett K, Cuncins AV, Hearn TC. Characterization of vertebral strength using digital radiographic analysis of bone structure. Med Phys 1995;22(5):611-615.
17. Rossman K. Spatial fluctuations of X-ray quanta and the recording of radiographic mottle. Am J Roentgenol 1963;90:863-869.
18. Webb S, ed. The physics of Medical Imaging. Bristol, Philadelphia: Adam Hilger, 1988.
19. Veenland JF, Grashuis JL, Van der Meer F, Beckers ALD, Gelsema ES. Estimation of fractal dimension in radiographs. Med Phys 1996;(4):585-594.
20. Morishita J, Doi K, Katsuragawa S, Monnier-Cholley L, MacMahon H. Computer-aided diagnosis for interstitial infiltrates in chest radiographs: optical-density dependence of texture measures. Med Phys 1995;22(9):1515-1522.
21. Brodatz P. Textures, a photographic album for artists & designers. (1st ed.) New York: Dover Publications, 1966.
22. Liu S-S, Jernigan ME. Texture analysis and discrimination in additive noise. CVGIP 1990;49:52-67.
23. Ohanian PP, Dubes RC. Performance evaluation for four classes of textural features. Pattern Recognition 1992;25(8):819-833.
24. Peitgen H-O, Saupe D. The science of fractal images. New York: Springer-Verlag, 1987.
25. Gonzalez RC, Woods RE. Digital Image Processing. Reading, Massachusetts: Addison-Wesley Publishing Company, 1992.
26. Van Bemmel JH, Musen MA, ed. Handbook of Medical Informatics. Springer, 1997.
27. Bencomo JA, Fallone BG. A logit model for the modulation transfer function of screen-film systems. Med Phys 1986;13(6):857-860.
28. Tou JT, Gonzalez RC. Pattern recognition principles. Reading, Massachusetts: Addison-Wesley Publishing Company, 1974.
29. Weszka JS, Dyer CR, Rosenfeld A. A comparative study of texture measures for terrain classification. IEEE Trans Syst Man Cybern 1976;6:269-285.
30. Conners RW, Harlow CA. A theoretical comparison of texture algorithms. IEEE Trans Pattern Anal Machine Intell 1980;PAMI-2(3):204-222.

Chapter 4

Estimation of fractal dimension in radiographs

J.F. Veenland[1,2], J.L. Grashuis[1,2], F. van der Meer[2], A.L.D. Beckers[1], E.S. Gelsema[1]

[1] Department of Medical Informatics,
[2] Department of Radiology, Erasmus University Rotterdam, The Netherlands

Abstract

In the last decade, the fractal dimension has become a popular parameter to characterize image textures. Also in radiographs, various procedures have been used to estimate the fractal dimension. However, certain characteristics of the radiographic process, e.g. noise and blurring, interfere with the straightforward application of these estimation methods.

In this study, the influence of quantum noise and image blur on several estimation methods was quantified by simulating the effect of quantum noise and the effect of Modulation Transfer Functions, corresponding with different screen-film combinations, on computer generated fractal images. The results are extrapolated to explain the effect of film-grain noise on fractal dimension estimation.

The effect of noise is that, irrespective of the noise source, the fractal dimension is overestimated, especially for lower fractal dimensions. On the other hand, blurring results in an underestimation of the dimensions. The effect of blurring is dependent on the estimation method used; the dimension estimates by the Power Spectrum Method are lowered with a constant value, whereas the underestimation by the methods working in the spatial domain is dependent on the given dimension.

The influence of the MTF and noise on fractal dimension estimation seriously limits the comparability of fractal dimensions estimated from radiographs which differ in noise content or MTF. Only when the Power Spectrum Method is used, it is possible to correct for the influence of different MTFs of screen-film combinations. It is concluded that only when using the same object-focus distance, the same exposure conditions, the same digitizer at the same resolution, can fractal dimensions as estimated in radiographs be reliably compared.

***Key words:** Fractal dimension, Texture Analysis, Radiographs, Modulation Transfer Function*

1. Introduction

The fractal dimension is a concept which has, since the work of Mandelbrot[1], become widely used in image analysis. Several authors have applied the fractal dimension for the characterization[2-5], or the segmentation[6,7] of structures, or for object recognition[8]. What makes the fractal dimension such an attractive image descriptor? One explanation is that the surfaces of many natural objects can be described by a fractal model. Under some restrictions, the fractal dimension of an object can be computed from its image: the intensity surface of the image reflects the fractal dimension of the 3-D surface of the object[9]. Using fractal models, natural looking surfaces can also be generated. Furthermore, a correspondence exists between the fractal dimension and the perceived roughness of a surface: a smooth surface corresponds to a dimension close to 2, as opposed to a rough surface which approximates a dimension of 3. A true fractal surface is characterized by independence of scale: on every scale the structure is similar. In an image intensity surface this scale independency must not only apply to the two spatial dimensions, but also to the intensity dimension.

Several methods have been developed to estimate the fractal dimension of an image. Since the methods differ in their definition of the scaling parameter and quantities measured, the estimated dimension of an image varies with the algorithm used[10,11]. Furthermore, in real images the characteristics of the imaging technique influence the information content of the image and thus the estimated dimension. As long as one estimation method is used within a single image, for example, for segmentation purposes, the estimated fractal dimensions of different regions can be readily compared. However, when the fractal dimension is used to characterize structures in different radiographs, the influence of the imaging process should be taken into account. In screen-film radiography, on which we will concentrate in this chapter, a cascade of degrading processes can be discerned in the imaging chain. By simulating these processes[12], the effect on image content and on image analysis methods can be studied.

In this chapter we will investigate the influence of certain radiographic imaging characteristics on various fractal dimension estimation methods. First, computer generated fractal images are used to obtain reference values for each fractal dimension estimation method. Subsequently, the effect of the Modulation Transfer Function of a screen-film system and the effect of quantum noise are simulated on the fractal images. The influence of both processes on the estimation of the fractal dimension by the various algorithms is quantified. These results are extrapolated to explain the effect of film-grain noise on fractal dimension estimation.

2. Methods for fractal dimension estimation

In literature, a wide variety of fractal dimension estimation methods has been described. All these methods are based on the same principle: an image characteristic is measured as a function of a scale parameter. The relation between these two quantities can be described by a straight line in a log-log domain. The slope, computed by linear regression analysis, is linearly related to the fractal dimension. The various estimation methods differ in their definition of the image characteristic and the scale parameter. For our study, we selected a number of methods which have been applied to radiographic images. Methods which do not meet the intensity scale invariance requirement were not considered[2]. Four methods were selected: the Power Spectrum Method, the Intensity Variance Method, the Variation Method, and the Blanket Method. The Power Spectrum Method and the Blanket Method were described extensively in Chapter 2. The Intensity Variance Method and the Variation Method are described below.

Intensity Variance Method
If the intensity surface of the image is considered to be generated by a fractal Brownian function, the behaviour of the second order statistics can be described by[9]

$$VAR(I(x + \Delta x) - I(x)) \propto (\Delta x)^{2H},\tag{1}$$

where H is the Hausdorff-Besicovitch dimension.

This proportionality describes one of the basic properties of a fractal Brownian function extended to an intensity surface. The differences in intensity $I(x)$, measured in two points separated by a distance Δx, have a Gaussian distribution. The relation between the variance of this distribution and the distance Δx is given in Eq. (1). Whereas Pentland[9] used this property to test the "fractalness" of images, this method can be employed in itself to compute the dimension H from the log variance versus the log distance plot. The Hausdorff-Besicovitch dimension is related to the fractal dimension D by

$$D = E + 1 - H\tag{2}$$

where E is the topological dimension of the surface.

Variation Method

This method described by Dubuc et al.[13] relates the maximum difference in intensity within a certain neighborhood of a pixel to the size of this neighborhood:

$$V(x, y, B) = \underset{(m,n) \in B}{MAX} \ I(m,n) - \underset{(m,n) \in B}{MIN} \ I(m,n) \qquad (3)$$

The neighborhood B of the pixel (x,y) can be parameterized (ε) by the radius, when the shape is circular, or by the edge length, in case of a square. The total variation, or volume $V(\varepsilon)$ is defined as the sum of all maximum differences, determined for each pixel in its neighborhood B. The dimension D can then be computed from the following relation

$$V(\varepsilon) \propto \varepsilon^{2-D} \qquad (4)$$

by fitting a straight line in the log $V(\varepsilon)$ versus log ε plot. It has been shown[11] that the fractal dimension thus computed deviates considerably from the given fractal dimension, especially for higher dimensions. The Blanket Method can be considered a modification of this method.

3. Material and Methods

Two models were used. In the first one, we modeled the X-ray intensity distribution in front of the screen as a fractal image. Since quantum noise and light photon scattering in the screen-film system have a considerable detrimental effect on fractal dimension estimation[14], we simulated these two processes: Poisson noise and two different Modulation Transfer Functions were modeled, respectively. The fractal dimension estimation methods were applied to the resulting, modeled, intensity distribution of the light in front of the film. In the second model, a fractal object is placed in front of the screen: the density distribution of the X-rayed object is modeled as a fractal image. Quantum noise and light photon scattering were simulated in the same way as in the first model. The response of the film was assumed to be linear. In addition, the effect of film-grain noise was simulated by adding Gaussian noise, and the effect of the Modulation Transfer Function of the digitizer was modeled. The characteristic curve of the digitizer was assumed to be linear. The fractal dimension estimation methods were applied to the resulting digital image.
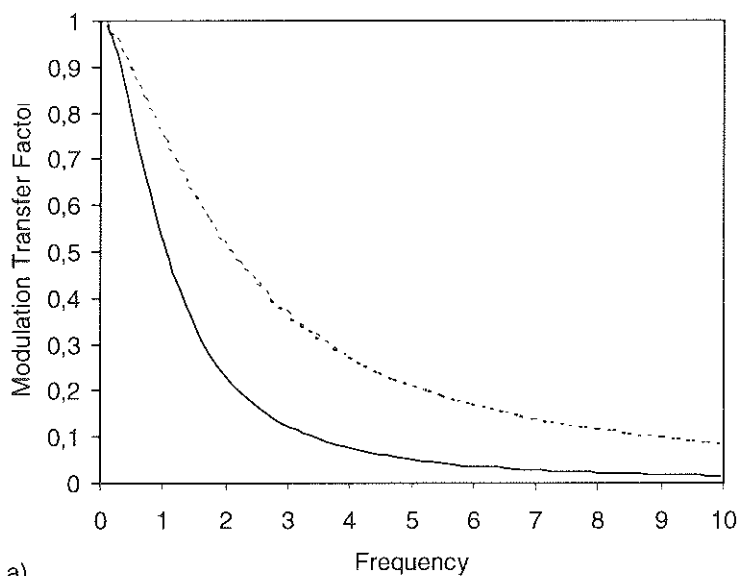
a)
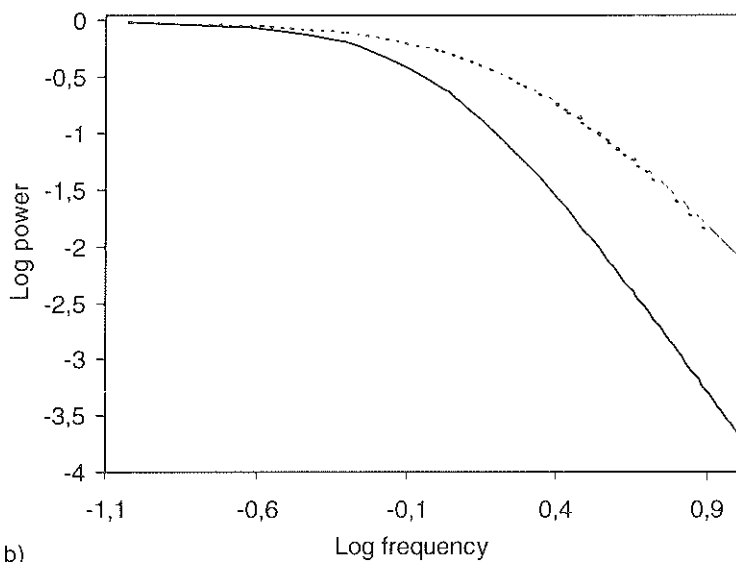
b)

Fig 1. *(a) Modulation Transfer Functions computed by the logit model and (b) the relation between the log Fourier Power and log frequency of the model of the two screen-film combinations used: Lanex Regular/0G (solid line) and Lanex Fine/0G (dashed line).*

Synthetic Fractal Images

Fractal images were generated using the spectral synthesis method[15]. By this method, Gaussian noise is added to the power spectrum corresponding to an ideal fractal with a given dimension, inverse Fourier transformation of this spectrum results in a fractal image. For each dimension, ranging from 2.1 to 2.9 in steps of 0.1, 10 images (256x256x8 bits) were synthesized.

Noise

The emission of quanta by an X-ray source, and therefore the absorption of these quanta in the intensifying screen is a stochastic process which can be modelled by a Poisson process. In our simulation experiments, the intensity values of the generated fractal images are considered to be proportional to the mean number of X-ray quanta absorbed in the screen-film system.

The intensity value of the fractal image was replaced by realizations drawn from a Poisson distribution, with mean value equal to the original intensity. Different signal to noise ratios are simulated by varying the noise level relative to the signal.

The film grain noise can be modelled as a zero-mean Gaussian process.

Modulation Transfer Function

The Modulation Transfer Function (MTF) can be used to describe the spatial resolution characteristics of screen-film systems. Whereas the lower spatial frequencies of the object spectrum projected on the radiograph are only slightly attenuated, the transfer of the higher frequencies is reduced. One of the causes is the scattering of light photons in the screen-film system. The influence of the scattering process can be spatially described by the point spread function (PSF). The modulus of the Fourier transform of the PSF is the modulation transfer function. We modelled the MTF using a logit function as described by Bencomo et al.[16].

$$logit\ MTF(f) = \log\{MTF(f)/(1 - MTF(f))\} \tag{5}$$

By logit analysis, the MTF of screen-film systems can be characterized by two parameters, a and b. These parameters describe the straight line which results when logit MTF is plotted against the log of the frequency f.

$$logit\ MTF(f) = b\log(f) + a \tag{6}$$

A complete analysis of the influence of the MTF on fractal dimension estimation would require the exploration of the two-dimensional (a,b)-space. Since our main

goal is to demonstrate the effect of the MTF, we used the parameters of two commonly used screen-film combinations, the Lanex Fine/0G and Lanex Regular/0G of Kodak, to model a relatively sharp and unsharp screen-film system, respectively. Figure 1(a) shows the MTF of the two screen-film combinations. The log power versus the log frequency for the two screen-film combinations is depicted in Fig. 1(b).

In our simulation experiments, we assumed that the absorption of one X-ray quantum results in the emission of a fixed number of light photons. The scattering process of these photons was simulated by convolving the synthetic images with the point spread functions of the different screen-film systems. This convolution was implemented in the Fourier domain using the transfer functions corresponding to the different MTFs, assuming a sampling distance of 50 μm. We simulated the effect of the MTF on clean fractal images and on images corrupted by quantum noise.

Linear Regression Analysis
As described in Chapter 2, the fractal dimension is computed by fitting a regression line in a plot of log image characteristic versus log scale parameter. It is well recognized that the fractal behaviour of natural images is limited to a certain range of scales. Therefore, the linear regression analysis should not take all data-points, deduced from all available resolution scales, into account: an upper and lower boundary must be specified for the scale parameter. Some authors determine the boundaries by visual inspection in each log-log plot. Since the effects of noise and MTF vary over the range, an objective comparison is only possible when fixed boundaries are chosen for each method. The range used for linear regression analysis is indicated by an arrow on the horizontal axis in Figures 2(a), 3(a), 4(a), 5(a), 6(a) and 7(a). To demonstrate the effect of MTF and noise more clearly, a rather large range is chosen for the linear regression.

Using the simulation components as described above, five types of experiments were performed.
A. *Comparison of the different methods.*
   First, for each generated fractal image and for each method, the fractal dimension was computed from the log-log plot using regression analysis as described. Then, for each group of 10 images for each given dimension, the mean fractal dimension and its standard deviation were computed. The H dimension, as estimated by the Power Spectrum Method and the Intensity Variance Method, was converted to the fractal dimension D using Eq. (2).

These data were used to compare the different methods, and to provide the reference data for the other experiments.

B. *Influence of the MTF.*
    Ideal fractal images were used to determine the influence of the Modulation Transfer Function alone on the different image characteristics as measured by the different estimation methods. Using linear regression analysis, the effect on each of the fractal dimension estimation methods was evaluated.

C. *Influence of quantum noise.*
    The ideal fractal images were corrupted by Poisson noise to quantify the influence of noise on the values of the image characteristics. The effect on the fractal dimensions, as estimated by different methods was evaluated. Different noise levels were simulated to illustrate the gradual effect of noise on dimension estimation.

D. *Combined influence of quantum noise and MTF.*
    In the actual radiographic process, both MTF and noise corrupt the image. The combined effect on the image characteristics was compared with the effects of MTF and noise separately. The influence of the combined effect on the fractal dimension estimates, as obtained by various methods, was quantified.

E. *Influence of quantum noise, MTF, and film-grain noise.*
    In addition to the effects of quantum noise and MTF, the influence of film-grain noise is simulated. Hereby, the limitations of correcting procedures, such as deconvolution, was demonstrated.

4. Results

A. Comparison of the different methods

Table 1. summarizes the fractal dimension as estimated by the different methods, averaged over 10 images for each chosen dimension. The estimates by the Power Spectrum Method differ only slightly from the given dimensions. This was to be expected since the generation procedure of the fractal images and the Power Spectrum Method are closely related. The estimates by the Intensity Variance Method and the Blanket Method deviate from the estimates by the Power Spectrum Method in that respect that the higher dimensions are estimated lower whereas the lower dimensions are given a higher estimate. The Variation Method results in the same ordering of the estimated dimensions, but the range is reduced to half of the range as estimated by the Power Spectrum Method.

Table 1. Average and standard deviation of the fractal dimension estimated by different methods on 10 fractal images for each given dimension D (PSM: Power Spectrum Method, IVM: Intensity Variance Method, VM: Variation Method, BM: Blanket Method).

	PSM		IVM		VM		BM	
D	Mean	SD	Mean	SD	Mean	SD	Mean	SD
2.1	2.108	0.034	2.185	0.018	2.152	0.022	2.240	0.047
2.2	2.220	0.039	2.242	0.033	2.194	0.020	2.296	0.044
2.3	2.318	0.028	2.314	0.014	2.250	0.011	2.388	0.025
2.4	2.406	0.022	2.374	0.012	2.286	0.013	2.431	0.037
2.5	2.491	0.024	2.448	0.014	2.335	0.012	2.515	0.039
2.6	2.622	0.030	2.529	0.012	2.395	0.007	2.619	0.015
2.7	2.717	0.065	2.593	0.015	2.437	0.009	2.693	0.032
2.8	2.811	0.054	2.654	0.012	2.477	0.011	2.771	0.047
2.9	2.931	0.037	2.716	0.011	2.515	0.004	2.847	0.016

B. Influence of the MTF

The effect of the MTF of the sharp screen-film system on the log power versus log frequency plot, as used by the Power Spectrum Method, is shown in Fig. 2(a). In this figure, the regression lines for two ideal fractal images, with dimension 2.2 and 2.8, respectively, are drawn. The effect of the MTF on log power, and thus on the fractal dimension estimates can be deduced. It can be seen that the reduction of the power is a non-linear function of the frequency. Performing the least squares fit in a region with frequencies higher than those contained in the region defined by the arrow, will result in a steeper regression line and therefore in a lower fractal dimension estimate: the dimension estimate is dependent on the range of frequencies chosen for the linear regression. When comparing the effect of the MTF on log power for the two images, it can be seen that the influence of the MTF on the power spectrum is independent of the given dimension. Therefore, as long as a fixed region for linear regression analysis is used, the effect of the MTF is that the dimension D is underestimated by the same amount for all images, as is shown in Fig. 2(b). This constant, which can be directly related to the MTF, is smaller for the sharper screen.

Table 2. *Influence of the MTF on the average and standard deviation of the fractal dimension, estimated by different methods on 10 fractal images for each given dimension D (PSM:Power Spectrum Method, IVM:Intensity Variance Method, VM:Variation Method, BM:Blanket Method).*

	PSM		IVM		VM		BM	
D	Avg	SD	Avg	SD	Avg	SD	Avg	SD
2.1	1.680	0.034	2.078	0.011	2.079	0.020	2.155	0.042
2.2	1.792	0.039	2.097	0.021	2.097	0.016	2.181	0.038
2.3	1.890	0.028	2.121	0.011	2.131	0.009	2.247	0.020
2.4	1.997	0.022	2.143	0.010	2.143	0.015	2.261	0.037
2.5	2.063	0.024	2.173	0.013	2.172	0.014	2.315	0.041
2.6	2.193	0.030	2.215	0.010	2.211	0.007	2.382	0.018
2.7	2.288	0.065	2.252	0.019	2.239	0.012	2.430	0.031
2.8	2.383	0.054	2.286	0.016	2.268	0.017	2.480	0.050
2.9	2.502	0.037	2.336	0.010	2.303	0.005	2.545	0.014

The effect of the MTF on the area, as computed by the Blanket Method, is illustrated in Fig. 3(a). The regression lines for two ideal fractal images are shown. The area is reduced, especially for the lower values of e, resulting in a less steep regression line and an underestimation of the fractal dimension as compared to the values obtained from ideal images, using the same method. Comparing the effect of the MTF for the two images with dimension 2.2 and 2.8, it can be seen that the area for images with a higher fractal dimension is relatively more reduced. This effect of the MTF, namely an underestimation of especially the higher fractal dimensions is displayed in Fig. 3(b) for the two screen-film combinations. The range of the estimated dimensions is drastically reduced.

The Intensity Variance Method and the Variation Method respond in the same way as the Blanket Method to the effect of the MTF. Table 2. summarizes the effect of the MTF corresponding to the Lanex Fine/0G screen-film system on the estimation of the fractal dimension by the different methods.
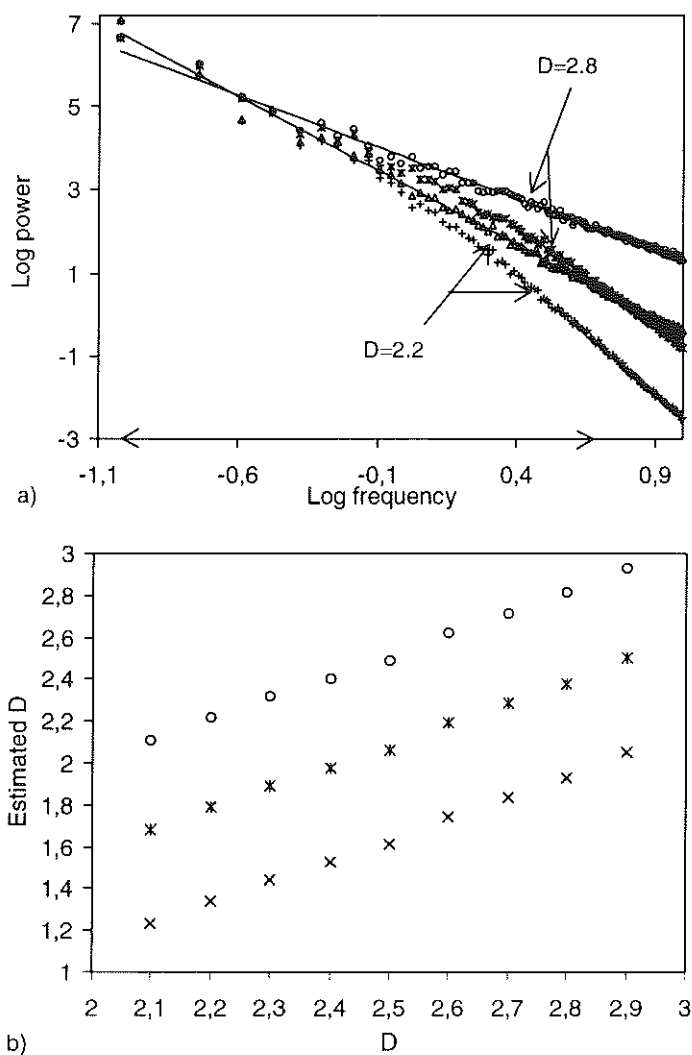
*Fig 2. The influence of the MTF on the Power Spectrum Method: a) the relation between log power and log frequency for ideal fractal images (Δ and o) and for images on which the MTF of Lanex Fine is simulated (+ and *), for images with fractal dimension 2.2 and 2.8, respectively; b) the dimension estimates as obtained from ideal images (o) and from images on which the effect of the MTFs of Lanex Fine (*) and Lanex Regular (×) is simulated.*
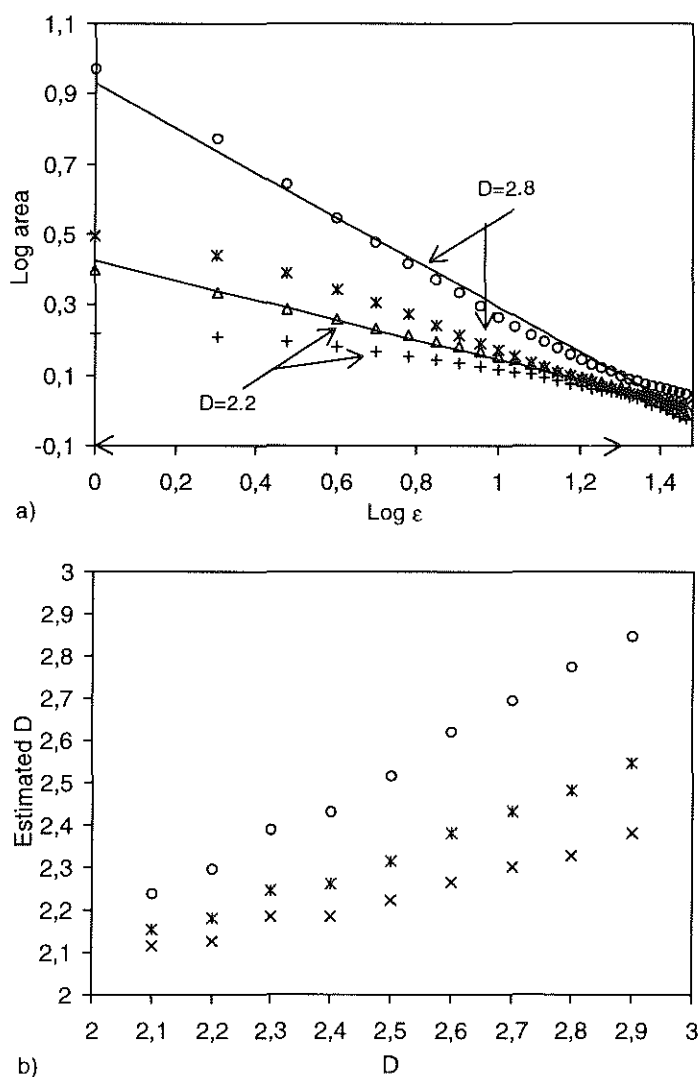
*Fig 3. The influence of the MTF on the Blanket Method: a) the relation between log area and log ε for ideal fractal images (Δ and o) and for images on which the MTF of Lanex Fine is simulated (+ and *), for images with fractal dimension 2.2 and 2.8, respectively; b) the dimension estimates as obtained from ideal images (o) and from images on which the effect of the MTFs of Lanex Fine (*) and Lanex Regular (×) is simulated.*
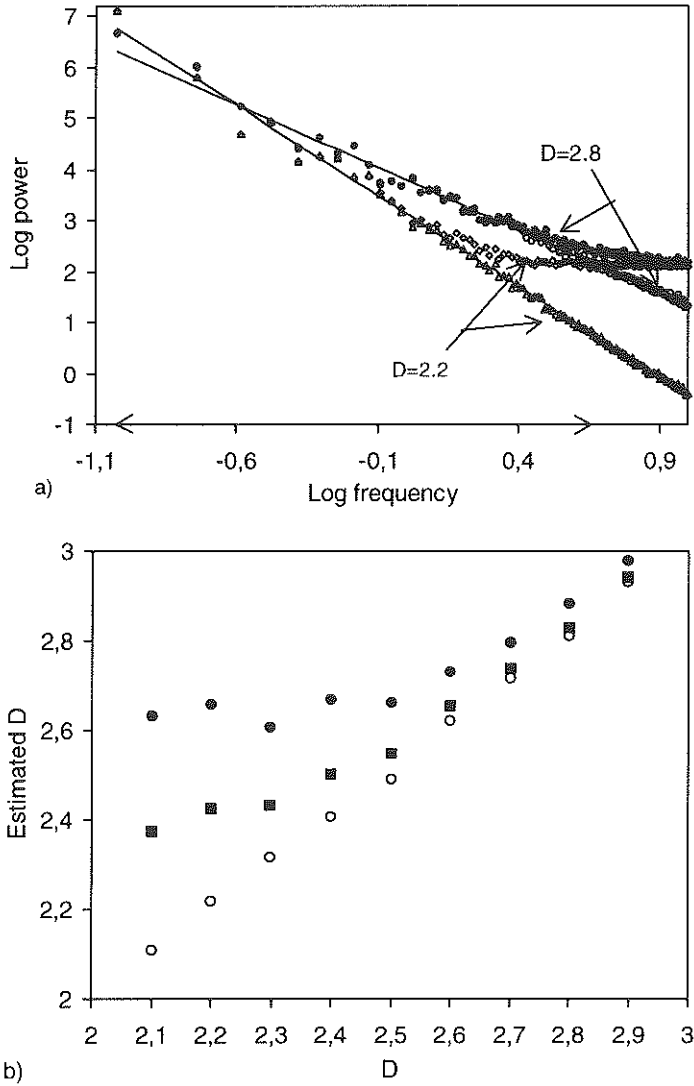
Fig 4. *The influence of Poisson noise on the Power Spectrum Method: a) the relation between log power and log frequency for ideal fractal images with dimensions 2.2 and 2.8 (Δ and ο), respectively and for their noisy versions (◊ and •); b) the results of the dimension estimation applied to ideal images (ο) and to images with noise levels 0.5 and 1 (▦ and •respectively)*

72

C. Influence of noise

The effect of Poisson noise on the power spectrum is shown in Fig. 4(a). Due to the flattening effect of the Poisson noise on the tail of the power spectrum, the slope of the regression line will be less steep, resulting in a higher dimension. Comparing the power spectra of the two images with given dimension 2.2 and 2.8, it can be seen that the influence of noise on the slope of the regression line is more prominent for the lower dimensions. This effect, which results in a drastic reduction of the range of the dimension estimates, is displayed for two noise levels in Fig. 4(b). Depicting the estimated dimension as a function of noise level (noise variance divided by mean signal value), it is demonstrated that the influence of noise on dimension estimation is gradual (Fig. 6(a)). The dimension estimates for the fractal image with given dimension 2.2 and 2.8 are not reliable above noise level 3 and 2, respectively, since the linearity conditions are violated ($r^2 < 0.90$). Choosing a smaller frequency range for linear regression analysis will lessen the influence of noise.

Figure 5(a) shows that as a result of the addition of noise, the area as measured by the Blanket Method, is enlarged especially for the lower values of e. This results in a steeper regression line: the dimension is overestimated. This overestimation due to Poisson noise is more prominent for the images with a lower fractal dimension, as illustrated in Fig. 5(b). The range of the dimension estimates is narrowed, dependent on the noise level. This effect is demonstrated in Fig. 6(b).

The Intensity Variance Method and the Variation Method respond in the same way as the Blanket Method to the addition of Poisson noise (not shown).

D. Combined influence of quantum noise and MTF

Figure 7(a) shows the combined effect of noise and MTF on the power spectrum and both effects separately for a fractal image with D=2.8. The effect of the combination of noise and MTF on the dimension estimation, displayed in Fig. 7(b), can be interpreted as the addition of two independent components.

This is not the case for the Blanket Method, where the effect of noise on the computed area is mitigated by the influence of the MTF (Fig. 8(a)). This dominant influence of the MTF on the dimension estimation is also shown in Fig. 8(b). The range of the estimated dimensions for the combined effects is smaller than for the MTF alone.

The Intensity Variance Method and the Variation Method respond in the same way as the Blanket Method to the combination of Poisson noise and MTF (not shown).

a)

b)
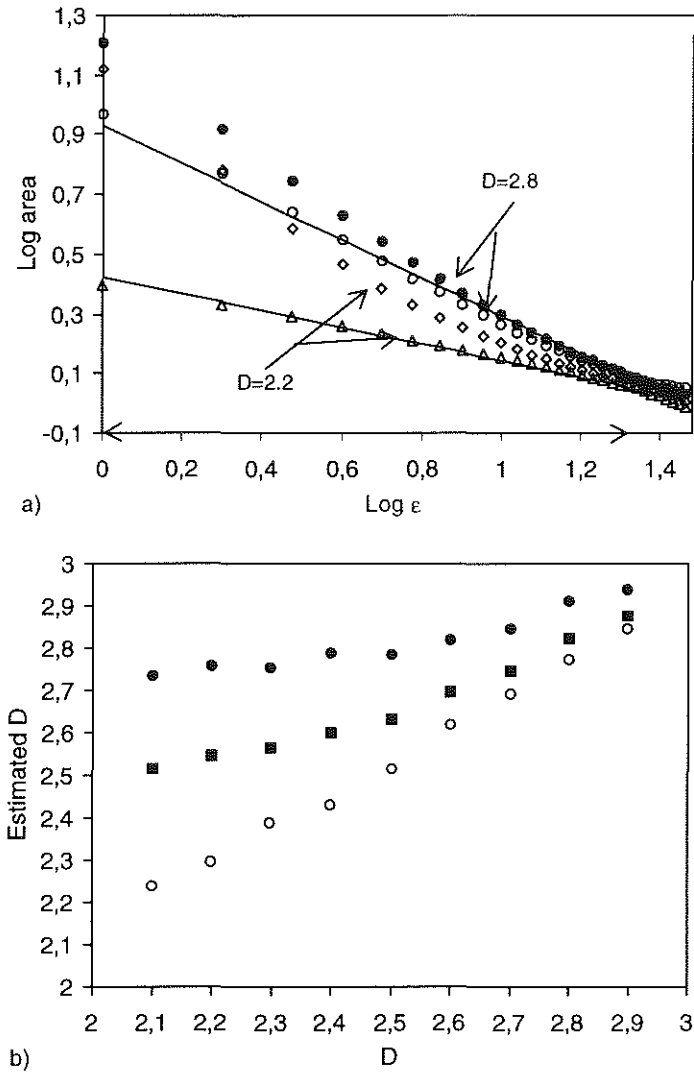
Fig 5. The influence of Poisson noise on the Blanket Method: a) the relation between log area and log ε for ideal fractal images with fractal dimension 2.2 and 2.8 (Δ and o) respectively, and for their noisy versions (◊ and •); b) the results of the dimension estimation applied to ideal images (o) and to images with noise levels 0.5 and 1 (■ and ●) respectively).

Fig 6. The influence of different noise levels on dimension estimation for images with fractal dimensions 2.2 and 2.8 (● and ◊ respectively) for the Power Spectrum Method (a) and the Blanket Method (b).

a)

b)
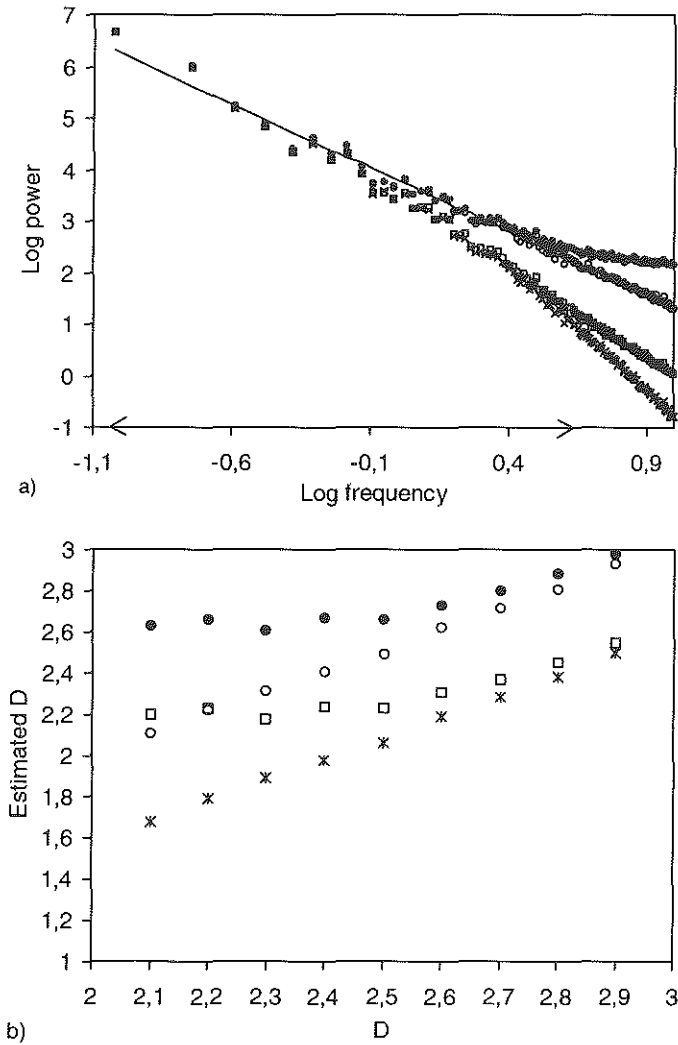
Fig 7. The influence of Poisson noise and the MTF on the Power Spectrum Method: (a) the relation between log power and log frequency for the ideal fractal image with D=2.8 (o), its noisy version (●), and the ideal and noisy image on which the MTF of Lanex Fine is simulated (* and □ respectively); (b) the dimension estimation applied to ideal images (o), noisy images (●), and ideal and noisy images on which the MTF of Lanex Fine is simulated (* and □ respectively).

76
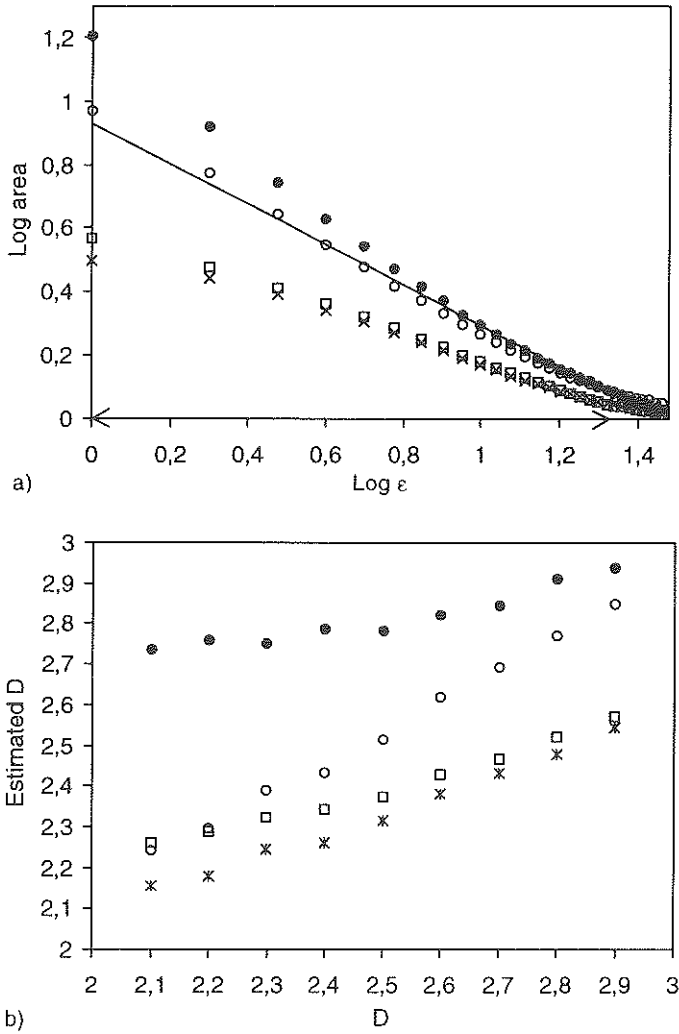
Fig 8. The influence of Poisson noise and the MTF on the Blanket Method: (a) the relation between log area and log ε for the ideal fractal image (○), the noisy image (●), and the ideal and noisy image on which the MTF of Lanex Fine is simulated (and □ respectively); (b) the dimension estimation applied to ideal images (○), noisy images (●), and the ideal and noisy image on which the MTF of Lanex Fine is simulated (* and □ respectively).*
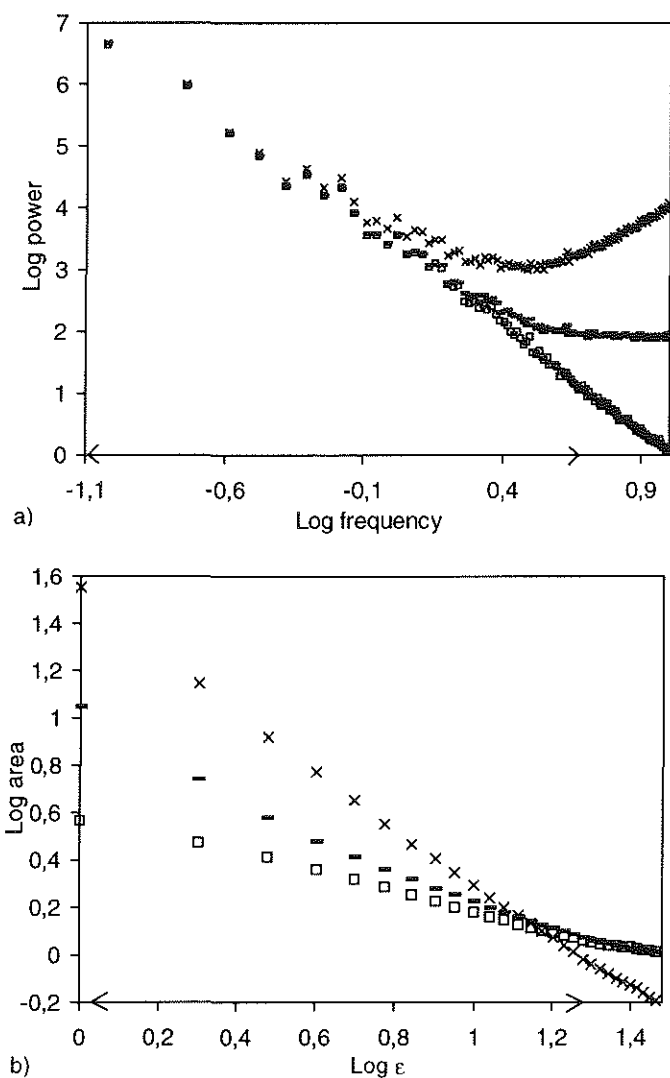
Fig 9. *The influence of film-grain noise and deconvolution on a) the Power Spectrum Method and b) the Blanket Method: the relation between respectively log power and log frequency and log area and log ε for the noisy fractal image (D=2.8) on which the MTF of Lanex Fine is simulated (□), after simulation of film-grain noise (−), and after deconvolution (x).*

E. Influence of quantum noise, MTF, and film-grain noise

In Figure 9 the additional effect of film-grain noise on the Power Spectrum and Blanket Method is shown. It can be seen that, for both methods, the effect of film-grain noise on dimension estimation is similar to the effect of quantum noise. Figure 9(a) also shows that the effect of deconvolution, in order to correct for the effect of the MTF, is limited due to the enhancement of noise.

5. Discussion

In the last decade, the fractal dimension has become very popular as a means to characterize image textures. To estimate the fractal dimension of an image, several methods measuring different fractal characteristics have been developed. In radiographs, fractal dimensions have been used for example to differentiate between osteoporotic and normal bone[3,17,18] and to characterize mammographic textures[5]. However, the radiographic process has certain characteristics which preclude a straightforward application of the fractal dimension estimation methods. In this chapter, we investigated, using simulation, the influence of two aspects of the radiographic imaging chain, namely quantum noise and the Modulation Transfer Function, on four different fractal dimension estimation methods. From these results the influence of film-grain noise and the influence of the MTF of the digitizer can be inferred.

A characteristic property of the fractal dimension of a fractal surface is its independence of scale in all three directions. So, the intensity scale in the image of a fractal surface can be varied independently of the spatial scales, without changing the fractal dimension. Therefore, methods which assume a fixed relation between the spatial and the intensity scale[2], were not considered here. Three dimension estimation methods which have been used to characterize textures in radiographs were selected: the Power Spectrum Method[7,17,18], the Blanket Method[6,19], and the Intensity Variance Method[6,20]. In addition, the Variation Method, which is related to the Blanket Method, was also evaluated.

Applying the different estimation methods to ideal fractal images it is seen that the various methods apparently estimate various aspects of "fractalness". This is in agreement with results of Sarkar et al.[10] and Huang et al.[11]. Compared to the Power Spectrum Method, all other methods, especially the Variation Method result in a reduced range of the dimension estimates, however, the ordering is preserved.

Using the dimension estimates as obtained by different methods from ideal fractal images as a reference for that method, the influence of the simulated

distortions was investigated. A distinction can be made between the Power Spectrum Method and the other three methods. Since the response of the Intensity Variance Method and the Variation Method can be deduced from the Blanket Method, only the effect of MTF and noise on the latter method will be discussed. Our simulation experiments showed indeed identical behaviour of the three methods.

Simulation of the effect of the MTF results in blurred images: the transfer of especially the higher frequencies of the ideal fractal image spectrum is reduced. The Power Spectrum Method estimates the dimension from the slope of the regression line, which can be drawn through the data points in the log power versus log frequency plot. Since the MTF influences the power spectra of images with different dimensions in the same way, the slopes of the regression lines are equally affected as long as the least squares fit is performed on the same frequency range, resulting in a systematic underestimation of the fractal dimension. This is confirmed by our results. The effect of the MTF on the dimension estimation by the Blanket Method is more intricate. The area, as measured by the Blanket Method, is reduced by the smoothing effect of the MTF. This effect is larger for the images with higher fractal dimensions which have rougher intensity surfaces. Since the range of the average estimates is reduced with respect to the Blanket Method estimates from ideal images, whereas the standard deviation for each given dimension is scarcely influenced (Table 2.), the ability to discriminate between images with different dimensions will be decreased. Comparing the effect of the MTF of two screen-film combinations, a relatively sharp and unsharp one (Figures 2(b) an 3(b)), it can be seen that the dimensions as estimated from radiographs obtained with different screen-film combinations cannot readily be compared. Only when the Power Spectrum Method is used, a correction for differences in MTF is feasible, provided the MTFs are known: the image spectrum is divided by the squared MTF. Chen et al.[14] estimate the MTF directly from a uniform region of the radiograph. It will be interesting to investigate whether the MTF, thus determined, is accurate enough to make a comparison between estimates from different screen-film combinations possible. Theoretically, deconvolution could remove the effect of the MTF. However, the noise present in the radiographs, e.g. film-grain noise, will be accentuated by such a deconvolution.

Poisson noise influences the different fractal dimension estimation methods in more or less the same way. For each method, the dimensions, especially the lower ones, are overestimated as compared to the reference values for that method. Fractal images with a higher fractal dimension have relatively more high frequencies in their power spectrum, resulting in a rougher appearance of the

intensity surface. Adding Poisson noise to a fractal image will increase the contribution of the high frequencies to the image power spectrum and will therefore make the intensity surface appear rougher, resulting in overestimation. This effect will be more pronounced for images with a relatively small high frequency content and thus with a smooth intensity surface. Some authors, applying the Power Spectrum Method, use visually determined upper frequency limits. This procedure limits the comparability between estimates from different images, as the contribution of noise and the MTF in the image spectrum is no longer comparable. When noise conditions and MTF are equivalent, a fixed frequency range, which excludes apparent noise, for different images should be used. The overall effect of noise is that the range of the estimated dimensions is reduced. Therefore, the ability to discriminate between images with different fractal dimensions is diminished.

The influence of film-grain noise can be inferred from the effect of quantum noise. By digitizing a radiograph, the influence of the MTF of the digitizer is incorporated in the power spectrum of the digitized image, and so exerts influence on the fractal dimension estimation. Magnifying the radiograph optically, e.g. when using a CCD camera for digitization, will result in a frequency shift in the power spectrum of the radiographs, when plotted logarithmically, whereas the modulation of the relative frequencies by the digitizer will be the same, independent of the magnification. Therefore, similar digitizing conditions are required to enable a comparison between fractal dimensions estimated from different radiographs.

Our simulation model was confined to the effect of noise and the MTF. We are well aware of the fact that we left out of consideration several factors which also influence fractal dimension estimation. Our aim was to illustrate the limiting effect, resulting from the MTF and noise, on the discriminative power of fractal dimension estimates and on the comparability of fractal dimensions estimated from different radiographs. Since these two processes have already a rather detrimental influence, other factors will only further deteriorate the estimates.

6. Conclusions

In this Chapter, the influences of the radiographic process on different fractal dimension estimation procedures have been made explicit using simulation experiments. The effects of the Modulation Transfer Function and noise on four different fractal estimation methods have been discussed. From the results of these experiments, distorting effects of other processes can be inferred.

It is seen that when fractal dimension estimates are used for the purpose of discriminating between different images, the effects of MTF and noise are detrimental: in most cases the range of the estimates decreases. An exception to this general rule is the effect of the MTF on fractal dimension estimates as obtained with the Power Spectrum Method; the estimates are decreased by an amount which depends on the MTF but not on the reference value. In practice this means that when comparing fractal dimension estimates obtained with the Power Spectrum Method from radiographic images taken with the same screen-film combination, no corrections for the influence of the screen-film MTF are necessary.

In general, however, the object-focus distance, the exposure conditions, the screen-film system, the digitizer, the optical magnification of the radiograph, influence the fractal dimension via MTF and noise. Therefore, fractal dimensions, as estimated by one method in different radiographs, can only be reliably compared when conditions for radiographing and digitizing are as similar as possible.

References

1.   Mandelbrot BB. The fractal geometry of nature.New York: W.H. Freeman, 1977.
2.   Caldwell CB, Stapleton SJ, Holdsworth DW, et al. Characterisation of mammographic parenchymal pattern by fractal dimension. Phys Med Biol 1990;35:235-247.
3.   Caligiuri P, Giger ML, Favus MJ. Multifractal radiographic analysis of osteoporosis. Med Phys 1994;21(4):503-508.
4.   Benhamou CL, Lespessailles E, Jacquet G, et al. Fractal organization of trabecular bone images on calcaneus radiographs. J Bone Miner Res 1994;9(12):1909-1918.
5.   Priebe CE, Solka JL, Lorey RA, et al. The application of fractal analysis to mammographic tissue classification. Cancer Letters 1994;77:183-189.
6.   Dellepiane S, Serpico SB, Vernazza G, Viviani R. Fractal-based image analysis in radiological applications. Proc SPIE 1987;845:396-403.
7.   Kuklinski WS, Chandra K, Ruttiman UE, Webber RL. Application of fractal texture analysis to segmentation of dental radiographs. Proc SPIE 1989;1092:111-117.
8.   Stein MC. Fractal image models and object detection. Proc SPIE 1987;845:293-300.
9.   Pentland AP. Fractal-based description of natural scenes. IEEE Trans Pattern Anal Machine Intell 1984;6:661-674.
10. Sarkar N, Chaudhuri BB. An efficient approach to estimate fractal dimension of textural images. Pattern Recognition 1992;25(9):1035-1041.
11. Huang Q, Lorch JR, Dubes RC. Can the fractal dimension of images be measured? Pattern Recognition 1994;27(3):339-349.

12. Van Metter R, Dillon PLP, Huff KE, Rabbani M. Computer simulation of radiographic screen-film images. Proc SPIE 1986;626:82-99.
13. Dubuc B, Roques-Carmes C, Tricot C, Zucker SW. The variation method: a technique to estimate the fractal dimension of surfaces. Proc SPIE 1987;845:241-248.
14. Chen J, Zheng B, Chang Y-H, Shaw CC, Towers JD, Gur D. Fractal analysis of trabecular patterns in projection radiographs. Invest Radiol 1994;29(6):624-629.
15. Peitgen H-O, Saupe D. The science of fractal images. New York: Springer-Verlag, 1987.
16. Bencomo JA, Fallone BG. A logit model for the modulation transfer function of screen-film systems. Med Phys 1986;13(6):857-860.
17. Ruttimann ES, Webber RL, Hazelrig JB. Fractal dimension from radiographs of peridental alveolar bone. Oral Surg Oral Med Oral Path 1992;74:98-110.
18. Webber RL, Underhill TE, Horton RA, Dixon RL, Pope TL. Predicting osseous changes in ankle fractures. IEEE Eng Med Biol 1993;12:103-110.
19. Lynch JA, Hawkes DJ, Buckland-Wright JC. Analysis of texture in macroradiographs of osteoarthritic knees using the fractal signature. Phys Med Biol 1991;36(6):709-722.
20. Chen C-C, Daponte JS, Fox MD. Fractal feature analysis and classification in medical imaging. IEEE Trans Med Imaging 1989;8(2):133-142.

Part II:

Post-mortem Studies

Chapter 5

Reproducibility of texture features in femur radiographs

J.F. Veenland[1,2], T.M. Link[3], J.L. Grashuis[1,2], E.S. Gelsema[1]

[1] Department of Medical Informatics
[2] Department of Radiology, Erasmus University Rotterdam, The Netherlands
[3] Department of Clinical Radiology, Münster, Germany

Abstract

Trabecular bone is a non-homogenous structure. Small variations in repositioning, inevitable in a clinical setting, may result in a slightly different texture visible in the radiograph. When using texture analysis methods to quantify the trabecular pattern in radiographs, these methods should be robust to texture variations caused by repositioning or small rotations. In this study we investigated the reproducibility of four texture analysis methods: the Spatial Gray-Level Dependence Method (SGLDM), the Fourier Method, fractal dimension estimation methods, and the Morphological Gradient Method.

For this study, the right femora of 10 females were used. Two experiments were performed: 1) four femora were repositioned each four times, and 2) ten femora were rotated over an angle of −15° and +15° with respect to the longitudinal axis. Direct Magnification (DIMA) radiographs with a 1.5 fold magnification were obtained for each repositioning and rotation. In four regions of interest in each femur, all texture features were determined. We defined reproducibility under repositioning as the percentage of the total variance of each feature which could be attributed to the variance between the subjects.

Results of the study indicate that most of the SGLDM and fractal features are well reproducible under repositioning. For the Fourier features and the morphology gradient features, a subset of features show a high reproducibility. The reproducibility under rotation depends on the location of the region of interest in the femur. The SGLDM and the morphology gradient method generate features which are the least influenced by rotation.

We conclude that the reproducibility of texture features is rather high, and that, therefore, texture analysis can be used for the analysis of the structure of clinical bone radiographs. Since the reproducibility of texture features may depend on the skeletal site and on the location of the region of interest, it is strongly advised to asses the reproducibility of texture features before evaluating the discriminative performance on clinical radiographs.

1. Introduction

The strength of bone is determined not only by density, but also by structure. Whereas for the assessment of bone density several sophisticated techniques have been developed, methods to measure bone structure in vivo are still in their infancy. Several authors have employed texture analysis methods to quantify the radiographic trabecular structure. They investigated the performance of different methods such as those based on fractal dimension[1-4], the Fourier power spectrum[5], and Sobel filters[6]. However, a radiograph displays a projection of a three-dimensional structure. When the structure is non-homogenous, as is the case with trabecular bone, rotations will have a profound effect on the texture visible in the radiograph. Whereas a texture analysis method may be expected to detect large changes in texture due to rotation, it should be robust for texture variations caused by small rotations. After all, the underlying 3-dimensional structure is not changed. In a clinical setting, such small rotations due to variations in positioning are inevitable. Furthermore, no two radiographs of one object will be exactly the same, even if the object is not repositioned. In each of the different components of the imaging chain small random variations may occur, which are reflected in the radiograph.

For bone density measurement methods, several studies have been performed to determine the accuracy and the reproducibility of the techniques. Where accuracy is defined as the degree of conformity of a measurement to a standard or true value, reproducibility can be defined as the degree of consistency among repeated measurements performed on the same object under conditions that are as close to identical as possible. For the determination of the accuracy a gold standard in the form of a standardized fantom is necessary. However, whereas in bone densitometry several phantoms have been developed, in the relatively young field of structure analysis such a gold standard is not yet available. Therefore, whereas it is not possible to determine the accuracy of a texture analysis method, the reproducibility can be determined. The differences to be measured by a method, between subjects or within one subject over a certain period, must be evaluated in relation to the reproducibility of this method. When the differences to be measured are of the same order of magnitude as the reproducibility, no meaning can be ascribed to differences found: the reproducibility functions as a lower bound for the detection of significant differences.

The aim of our study was to evaluate the reproducibility of different texture features generated by four selected texture analysis methods.
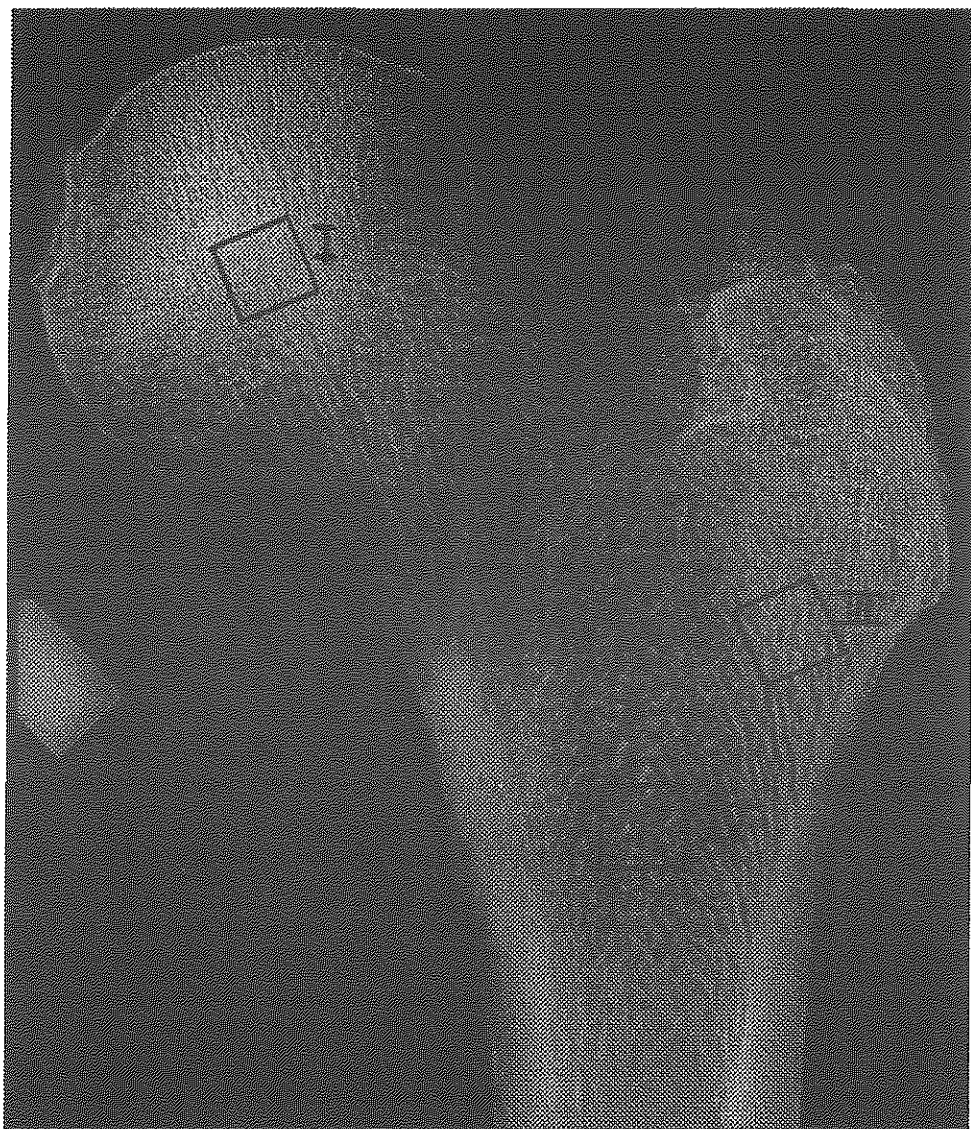
Fig 1. Magnification radiograph of a femur with the four selected regions.

2. Material and Methods

Material

We obtained the right femora from 10 females. The average age of the women was 80.9 years, ranging from 54.3 to 98.7 years. The femora were exarticulated in the hip and excised about two hand-widths above the knee, and after removal of surrounding soft tissue kept in a phenol/formaldehyde solution. The femora had no obvious pathology. The femora were placed in a specially designed holder. Direct Magnification (DIMA) radiographs of the femora were obtained using a microfocal X-ray tube (Feinfocus DIMA tec-013, Feinfocus Medical Systems, Garbsen, Germany). The images were taken with 1.5-fold magnification, a focus-film distance of 1.10 m, a focal spot size of 0.03–0.2 mm^2, and 60 kV. Computed radiography was used as an imaging system employing the storage phosphor technique (FCR-7000, Fuji, Tokyo, Japan). The cassette contained 1760x2140 elements with an element size of 200 µm x 200 µm.

In the resulting radiograph, four regions of interest were selected for each femur. In Figure 1 a DIMA radiograph of a femur with the selected regions is shown. Region I is located centrally in the femurhead, parallel to the principal direction of the trabecular structure. Region II, located in the upper part of the neck, and region III, located below the trochanter major, both contain parts of the principal tensile trabeculae[7]. Region IV is located centrally in the shaft. These regions, with a size of 1.0 x 1.0 cm^2, were digitized using a CCD-camera (256x256 pixels).

Texture analysis methods

In this study, we used four different texture analysis methods: the Spatial Gray-Level Dependence Method[8], the Fourier Method, two fractal dimension estimation methods (the Power Spectrum Method and the Blanket Method), and the Morphological Gradient Method. These methods are described extensively in Chapter 2. In the following text we will refer to the features generated by the different methods as the co-occurrence features, the Fourier features, the fractal features and the morphological gradient features, respectively. For each method, the features which are closely related according to their definition are clustered into feature sets. The resulting sets with constituting features are listed in Table 1.
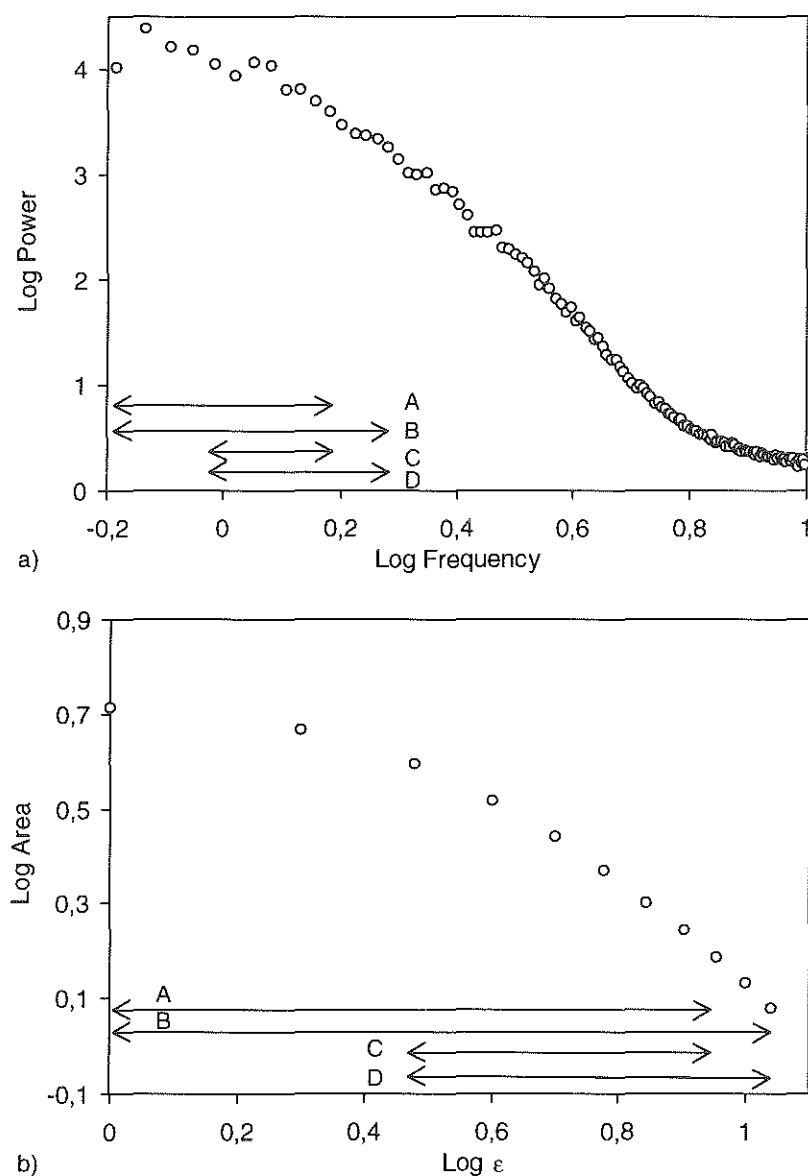
Fig 2. Log-log plots as used for the fractal dimension estimation by the Power Spectrum Method (a) and the Blanket Method (b). The ranges we used in linear regression analysis to obtain the four estimates are indicated at the bottom of the figures.

Co-occurrence features

We selected five co-occurrence features, namely Angular Second Moment, Contrast, Correlation, First Diagonal Moment, and Entropy, which are determined over three distances, (d=1, d=2, and d=4 pixels) and in two directions (θ=0°, and θ=90°). The direction 0° is chosen in line with the scan direction of the CCD-camera. This applies also for the other methods.

Fourier features

The group of Fourier features is subdivided into ten subsets. The first two sets encompass the oldest and most widely used spectrum features described by Bajcsy[9], summing the power of the Fourier Spectrum $P(r,\Theta)$ over radii (set 1) or angles (set 2). The third set consists of two features introduced by Katsuragawa[10]: the root-mean-square variations (*RMS*), and the first moment (*FMO*) of the Power spectrum. The Fourier features defined by Liu[11] are subdivided into seven sets, as listed in Table 1.

Fractal features

For the fractal features, two different fractal dimension estimation methods are used: the Power Spectrum Method[12] and the Blanket Method[13], modified after[1]. Both methods are based on the same principle: an image characteristic is measured as a function of a scale parameter. The relation between these two quantities can be described by a straight line in a log-log domain. The slope, computed by linear regression analysis, is linearly related to the fractal dimension. The two estimation methods differ in their definition of the image characteristic and the scale parameter, as described in Chapter 2. For each method, we visually selected four different ranges for the regression analysis. These ranges are illustrated in Fig. 2 a) and 2 b) for the Power Spectrum Method and the Blanket Method, respectively. Both methods are used on the image as a whole, and on the lines constituting the image, namely the lines parallel with the scan direction of the CCD-camera (x) and those perpendicular to that direction (y).

Morphological gradient features

The fourth method, the Morphological Gradient Method, was developed by ourselves especially for use in radiographs. The features are based on combinations of the basic operations of Mathematical Morphology[14]. Two morphological gradients are used: the G-gradient and the H-gradient. A complete description is given in Chapter 2.

Table 1. *Texture features generated by the different texture analysis methods.*

Method	Set number	Features
SGLDM	1	ASM, CON, COR, DMO1, ENT for d=1 θ=0°
	2	ASM, CON, COR, DMO1, ENT for d=2 θ=0°
	3	ASM, CON, COR, DMO1, ENT for d=4 θ=0°
	4	ASM, CON, COR, DMO1, ENT for d=1 θ=90°
	5	ASM, CON, COR, DMO1, ENT for d=2 θ=90°
	6	ASM, CON, COR, DMO1, ENT for d=4 θ=90°
FM	1	Radial features: $P_{1,2}$, $P_{2,4}$, $P_{4,8}$, $P_{8,16}$
	2	Angle features: $P_{0°,45°}$, $P_{45°,90°}$, $P_{90°,135°}$, $P_{135°,180°}$
	3	RMS, FMO
	4	Energy in major peak: f_1
	5	Laplacian of major and secondary peak: f_2, f_3
	6	Localization major and secondary peak: f_5, f_6, f_9, f_{10}, f_{11}, f_{12}, f_{13}
	7	Isotropy: f_7
	8	Moment of inertia: f_{17}, f_{18}, f_{19}
	9	Percentage energy in quadrant: f_{20}, f_{21}
	10	Relative entropy: f_{25}, f_{26}, f_{27}, f_{28}
FD	1	PSM, range A, B, C, D
	2	PSMX, range A, B, C, D
	3	PSMY, range A, B, C, D
	4	BM, range A, B, C, D
	5	BMX, range A, B, C, D
	6	BMY, range A, B, C, D
MGM	1	GR_5, GR_7, GR_9, GR_{11}
	2	$GSS_{7,3}$, $GSS_{9,5}$, $GSS_{11,7}$
	3	$GSR_{7,3}$, $GSR_{9,5}$, $GSR_{11,7}$
	4	$GO_{7,1}$, $GO_{9,3}$, $GO_{11,3}$, $GO_{11,5}$, $GO_{13,5}$
	5	HR_5, HR_7, HR_9, HR_{11}
	6	$HSS_{7,3}$, $HSS_{9,5}$, $HSS_{11,7}$
	7	$HSR_{7,3}$, $HSR_{9,5}$, $HSR_{11,7}$
	8	$HO_{7,1}$, $HO_{9,3}$, $HO_{11,3}$, $HO_{11,5}$, $HO_{13,5}$

Texture is constituted by the coherence, not the absolute values, of gray-levels. Therefore texture features should be independent of first order statistics as the average gray-level and the variance of the gray-levels. An ideal texture analysis method is invariant under strict monotonic gray-level transformations (requirement 1 in Chapter 1), but except for the Spatial Gray-Level Dependence Method none of the texture analysis methods meets this requirement. Therefore, the requirement is relaxed to "invariance under linear transformations of the gray-levels". All methods used, except the Fourier method, intrinsically fulfil this requirement. In order to be able to assess the reproducibility of the Fourier features, the images are pre-processed in such a way that the resulting images have equal average gray-values and equal variances of gray-levels.

In the Spatial Gray-Level Dependence Method, the first step consists of a histogram equalization in which the number of gray-levels is reduced. Thereafter, the co-occurrence features are determined at small scales of 1 to 4 pixels. However, histogram equalization has one drawback: when a background trend is present, this trend is enhanced at the expense of the more subtle gray-level variations at a smaller, spatial, scale. Performing a background correction ensures that, instead of the background trend, the smaller gray-level variations are quantified. We performed a simple background correction by using a large uniform filter (61x61 pixels).

Experiments
Two experiments were performed in order to evaluate the influence of reposition-ing and small rotations on the reproducibility of texture analysis features. In order to be able to reposition the femora as accurately as possible, the femora were provided with a resin foot, which fitted in a specially designed holder.

A. *Influence of repositioning.*
   Four femora were repositioned each four times. The repositioning encompassed replacing the femur in the holder and replacing the holder on the table. After each repositioning one radiograph was taken.

B. *Influence of rotation.*
   Ten femora were rotated over an angle of $-15°$ and $+15°$ with respect to the longitudinal axis. A radiograph was taken in each of the three positions.

Statistical Analysis

Two sources of variance in texture features are distinguished, i.e., the variance due to repositioning, V_P, and the variance present between different subjects, V_S.

The variance due to repositioning is computed from the radiographs of experiment 1: four femora are repositioned four times ($p=1..4$). The variance of the features f over the four radiographs is averaged for the four femora ($s=1..4$):

$$V_P = \frac{1}{4} \sum_{s=1}^{4} VAR \{f(s, p)\} \tag{1}$$

The variance in texture features which is present between different subjects is computed from the four radiographs of four femora (experiment 1), by the following procedure. Firstly, the texture features are averaged over the four radiographs of each femur, then the variance between these average features over the four femora is computed:

$$V_{SP} = VAR \left(\frac{1}{4} \sum_{p=1}^{4} f(s, p) \right) \tag{2}$$

The resulting variance, V_{SP}, combines the two sources of variance. The variance between subjects can be derived by subtracting the variance due to repositioning.

$$V_S = V_{SP} - \frac{V_P}{4} \tag{3}$$

The contribution of each source of variance is given as a percentage of the total variance present. The reproducibility is defined in this study as the fraction of the total variance which can be attributed to the variance between subjects. When the subject variance constitutes a large portion of the total variance present, then the contribution of the variance due to repositioning is small: the reproducibility is large. Since the reproducibility of a texture feature can be dependent on the underlying texture, we choose four regions to compute the reproducibility for each feature. The results for the four regions are combined into one output value using a scoring scheme, as described in Table 2. We discern four classes of reproducibility: 1. Excellent, 2. Good, 3. Moderate, and 4. Poor.

Table 2. *Classification scheme for texture features based on the contribution of the subject variance to the total variance.*

Number	Score	Criteria
1	Excellent	Subject variance is $\geq 85\%$ for at least three regions
2	Good	Subject variance is $\geq 75\%$ and $< 85\%$ for at least three regions
3	Moderate	Subject variance is $\geq 60\%$ and $< 75\%$ for at least three regions
4	Poor	Subject variance is $< 60\%$ for at least two regions

Using the radiographs taken in experiment 2, the variance in texture features due to rotation can be computed. In this part of the study, only the feature sets of which all members show a good or excellent reproducibility in experiment 1 are evaluated. For the sets of features thus selected, the variance in features over the three positions of the femur ($-15°$, $0°$, $+15°$) is computed for each femur separately. Then the average variance over the 10 femora is expressed as a percentage of the variance between the femora in the central position. Three classes of dependence on rotation are discerned: moderately independent (class 1), moderately dependent (class 2), and dependent (class 3), corresponding with respectively an average variance below 50%, between 50% and 100%, and over 100%. Since the effect of rotation depends on the location of the region of interest, this dependence of texture features on rotation is computed for each region separately.

3. Results

A. Influence of repositioning
In the Figures 3 a) through d), for each of the four texture analysis methods the distribution of the variance over the two sources of variance, namely repositioning, and subject, is shown for region I. The distribution of the variance for the regions II through IV follow a roughly similar pattern, therefore these are not shown separately. The reproducibility of the features, determined using the scheme in Table 2, is summarized in Table 3.

a)

b)
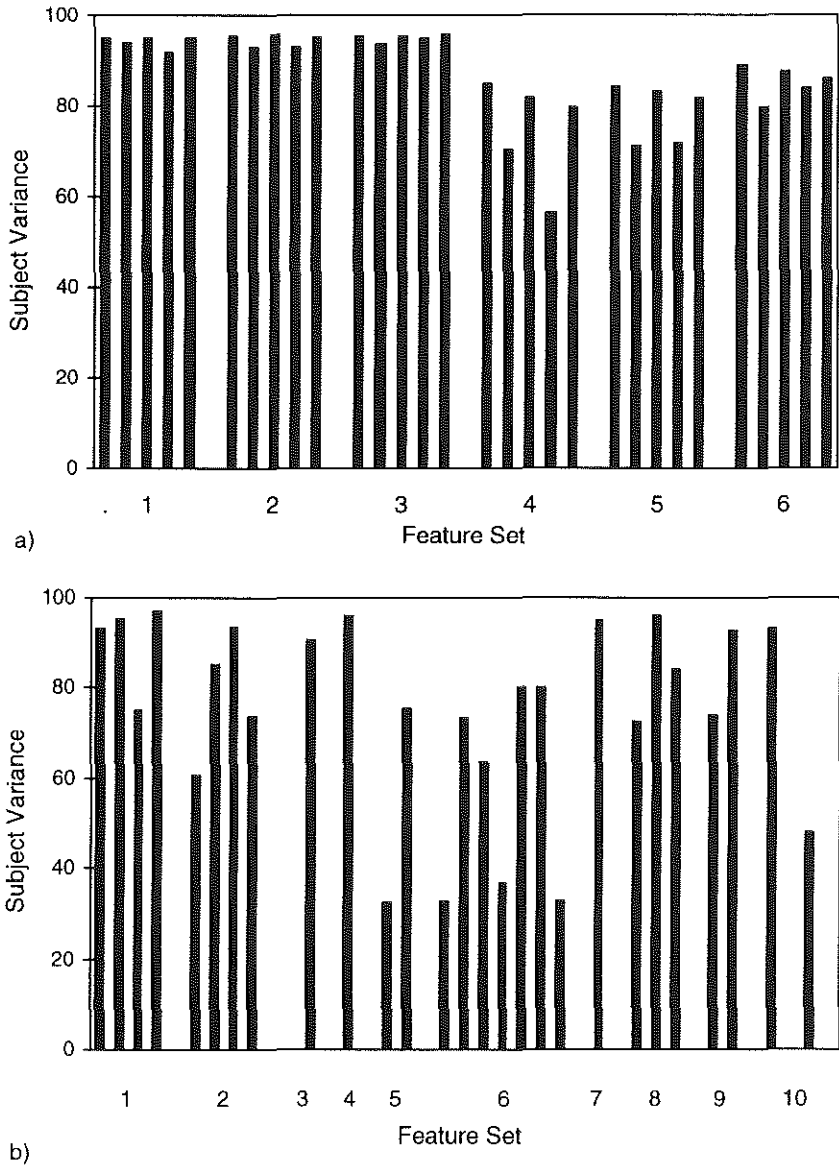
Fig 3. *The contribution of the subject variance to the total variance (in %) for the features of the four texture analysis methods, as determined in region 1: a) the Spatial Gray-Level Dependence Method, b) the Fourier Method, c) the Fractal dimension estimation methods, and d) the Morphological Gradient Method.*

c)

d)

99
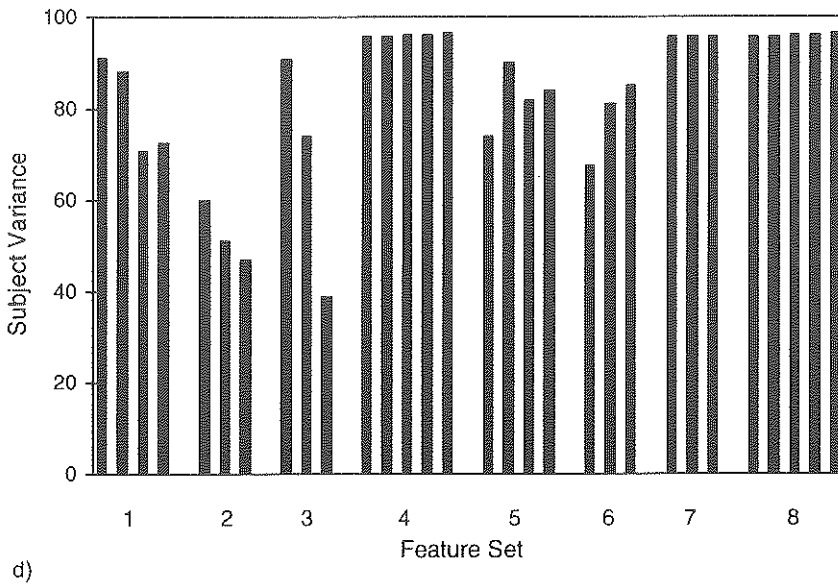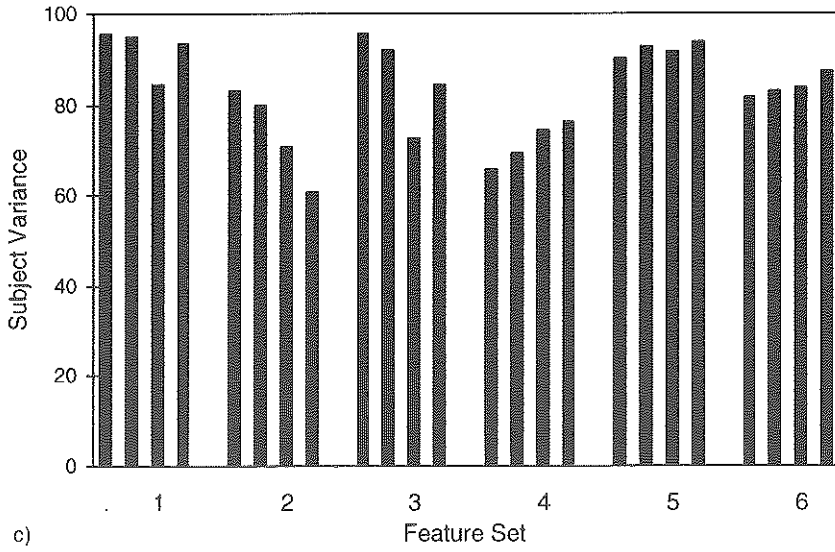
Co-occurrence features
Figure 3 a) shows that the reproducibility of the co-occurrence features in region I is very high, especially in direction 0° (sets 1 through 3). For the direction 90° (sets 4 through 6) the reproducibility varies with the feature used and the distance for which the feature is computed. On the average, the features determined using the larger distances (d=2, d=4) are somewhat better reproducible (sets 5 and 6). Furthermore, the features Angular Second Moment, Correlation, and Entropy (respectively, feature 1, 3, and 5 in each set) have a better reproducibility than the features Contrast and First Diagonal Moment (respectively, feature 2 and 4).

In Table 3, where the results for the different regions are summarized, the same pattern can be seen. The reproducibility of the features computed for the direction 0° is excellent, whereas at 90° the reproducibility of the features Contrast and First Diagonal Moment is moderate only for distance d=1 (set 4), for the other distances it is good to excellent.

Fourier features
For the Fourier features, the distribution of variance over the two sources for region I is shown in Figure 3 b). It should be noted that the Fourier features are on the whole rather sensitive to repositioning, especially the locally defined features as defined in[11] (sets 5 and 6). In Table 3 it can be seen that the rather poor reproducibility of these local spectrum features is not limited to region I. The amount of energy in the major peak (set 4) is the only local spectrum feature with a good reproducibility. The features of sets 7 through 9 show a good reproducibility in region I, however only the features of set 9 perform as well in at least two of the other regions. The last three features of set 10 are poorly reproducible. The radial Fourier features (set 1) are categorized as good to excellent, whereas the angular features (set 2) are, as a group, less well reproducible. The reproducibility of the root-mean-square variations (*RMS*, first feature of set 3) cannot be determined: all the radiographs have the same *RMS* value. This was to be expected, since the *RMS* value depends only on the average gray-value and the variance of gray-levels, which are set to identical values in the pre-processing step.

Table 3. *Classification of the features of the different texture analysis methods into four grades according to their reproducibility. The features are listed as set number: serial number in the set.*

Method	SGLDM	FM	FD	MGM
Features	Co-occurrence	Fourier	Fractal	Morphological Gradient
1 Excellent	1:1-5 2:1-5 3:1-5 6:1,3-5	1:1-3 2:2 9:1	1:1,2,4 2:2 5:1-4 6:4	1:1,2 3:1 4:1-5 7:1-3 8:1-5
2 Good	4:1,3,5 5:1-5 6:2	1:4 2:1 3:2 4:1 8:2,3 9:2 10:1	1:3 2:1,3,4 3:1,2,4 4:4 6:1-3	2:3 3:3
3 Moderate	4:2,4	2:4 6:4 8:1	3:3 4:1-3	1:3,4 2:2 3:2 5:2-4 6:3
4 Poor		2:3 5:1,2 6:1-3,5-7 7:1 10:2-4		2:1 5:1 6:1,2

Fractal features

Figure 3 c) shows that in region I the reproducibility of all fractal features is rather high. Using the Power Spectrum Method (sets 1 through 3), the fractal dimensions show a slightly lower reproducibility when determined on the x-lines (set 2). In contrast, using the Blanket Method (sets 4 through 6), the fractal dimension displays the lowest reproducibility when both directions are combined (set 4). Only the latter effect is also present in the other regions, therefore in Table 3 the

reproducibility of three of the features of set 4 is classified as moderate, whereas the reproducibility of the features of set 2 are classified as good or excellent.

Morphological Gradient features

The reproducibility of the morphological gradient features, in region I, is depicted in Figure 3 d). Three groups of features have a high reproducibility: the directional features for both gradients (sets 4 and 8) and the *HSR* features (set 7). In Table 3 the reproducibility of all these features is classified as excellent or good. On the other hand, the other H-gradient features (sets 5 and 6) are all classified as moderately or poorly reproducible.

B. Influence of rotation

Using the results of the repositioning experiment, as described above, for each method reproducible sets are selected. Only when the reproducibility of all the features of a set is good or excellent, a set is designated as reproducible. For example, of the sets of the Spatial Gray-Level Dependence Method, only one set is left out since some of the features show an only moderate reproducibility. The resulting sets of features are listed in Table 4. For these sets of features, the dependence on rotation was investigated for each region separately and classified as 1 if moderately independent, 2 if moderately dependent, and 3 if dependent. It can be seen that rotation has a major influence on most texture features determined in region III. Some of the Fourier features and the fractal features as determined in region IV by the Power Spectrum Method are fairly dependent on rotation. In region I and II, the co-occurrence features and the morphological gradient features show the least dependence on rotation.

Table 4. Classification of the features of the different texture analysis methods into three grades according to their dependence on rotation of the femur. The features are listed as set number: serial number in the set.

Method	Feature Set	Region 1	Region 2	Region 3	Region 4
SGLDM	1	1	1	1	1
	2	1	1	1	1
	3	1	1	1	1
	5	1	1	2	1
	6	1	1	2	1
FM	1	1	1	3	3
	3:2	2	1	2	3
	4	1	1	2	3
	9	2	2	3	1
	10:1	1	1	2	1
FD	1	2	1	2	3
	2	2	2	2	2
	5	1	1	1	1
	6	2	1	2	1
MGM	4	1	1	3	1
	7	1	1	1	2
	8	1	1	3	1

4. Discussion

A. Influence of repositioning

In the past, the trabecular structure, as projected in radiographs, has been quantified using different texture analysis methods based on, e.g., fractal dimension[1-4], the Fourier power spectrum[5], and Sobel filters[6]. However, trabecular bone is a non-homogenous three-dimensional structure. Therefore, the projection on a radiograph will vary with slight rotations, which are inevitable when repositioning. For a texture analysis method to be clinically useful, it should be robust to texture variations caused by small rotations. In this study we compare the reproducibility of texture features of four different texture analysis methods: the Spatial Gray-

Level Dependence Method, the Fourier Method, the fractal dimension (as estimated by the Power Spectrum Method and the Blanket Method), and the Morphological Gradient Method. By lack of a gold standard, the accuracy cannot be determined. The reproducibility is determined by repositioning and reradiographing a number of femora several times under conditions which are as similar as possible.

In each femur, four regions of interest were selected. The regions of interest were chosen in such a way that they contained characteristic parts of the trabecular trajectories in the femur. Furthermore, we made sure that the regions chosen could easily be located in different femora. The delineation of the regions was done manually, so in our study small variations in localization were inevitable. Automation of this step, using landmarks, may reduce the variation we found for repositioning. In the regions we selected, the main orientation of the trabecular structures differs. Furthermore, the variation in orientation of the structures varies. However, the reproducibility of texture features may be dependent on the orientation or the variation in orientation of the structures. Therefore in our study we developed a scheme, combining the results of the different regions. When the reproducibility of a texture feature was higher than a preset threshold, in at least three of the four regions, then the feature was rated accordingly. In this way, we allowed some tolerance for sensitivity of features for orientation. On the other hand, a texture feature which is only reproducible when the structures are oriented in a prescribed way, can be considered as not reproducible.

In this study, we chose to evaluate the influence of repositioning and rotation on texture features, since both processes will, inevitably, occur in a clinical context. These processes limit the applicability of texture analysis in clinical settings: texture features which are not reproducible under repositioning, can never be used in clinical practice. We did not evaluate the effect of redigitizing, region of interest selection, and reradiographing separately, since the effect of these processes on reproducibility are observed to be of a lower order and are contained in the effect of repositioning.

Co-occurrence features
Of the co-occurrence features, the features which are computed in line with the scan direction of the CCD-camera show an excellent reproducibility. The reproducibility of the features which are computed in the direction perpendicular to the scan direction varies with the definition of the feature and the distance for which the feature is computed. For the smallest distance the reproducibility of two features is only moderate, for the larger distances the reproducibility of all features is good to excellent.

Fourier features

The reproducibility of the Fourier features varies with the type of feature studied. The radial features show a good reproducibility, whereas the angular features are as a group not reproducible: the reproducibility of the features varies from poor to excellent. These angular features are sensitive to variations in orientation of the structures due to repositioning and region of interest selection. The same applies to the local and global spectrum features defined by Liu[11], with the exception of the amount of energy in the major peak, the percentage of energy in the first and second quadrant, and the relative entropy. The *FMO* defined by Katsuragawa[10] also shows a good reproducibility.

Fractal features

The reproducibility of all fractal features is rather high. The fractal dimension, as estimated by the Power Spectrum Method is highly reproducible, as long as it is determined on the whole image or on the x-lines constituting the image. In contrast, the fractal dimension, as estimated by the Blanket Method shows the best reproducibility when estimated on the x- or y-lines of the image. When determined on the whole image, the reproducibility is only moderate.

Morphological Gradient features

Of the morphological gradient features, three groups show an excellent reproducibility: the directional features for both types of gradients used, and the *HSR*-features. The *HSR*-features are sensitive to the size of structures in the image.

Influence of the scanning device

For the co-occurrence features, a difference in reproducibility for the two directions in which the features are generated was observed. Whereas the reproducibility of the features computed in line with the scan direction of the CCD-camera is excellent, the features computed perpendicular to the scan direction are less reproducible. Since the structures in the different regions are oriented differently with respect to the scan direction of the CCD-camera, it is not plausible that the orientation of the structures is the cause of this difference in reproducibility. It seems more likely that this difference originates from the different behavior of the CCD-camera in these two directions. The same effect was seen with the fractal dimension estimated by the Power Spectrum Method: the fractal dimension estimated from the x-lines of the image is more reproducible than the dimension estimated using the y-lines of the image. For the other texture methods used in this study, the Fourier Method and the Morphological Gradient Method, this effect was not perceptible since these methods do not contain features which are computed in

105

only one direction of the image. However, it should be noted that the scanning device has a profound influence on the texture content of the digitized image[15]. Therefore, it is advisable when comparing texture features in different radiographs, that the scanning conditions should be kept as similar as possible.

B. Influence of rotation

Due to repositioning, two type of rotations of the femur occur: in the plane of the radiograph, and perpendicular to the plane of the radiograph. The result of the first rotation is a different orientation of the structures: a rotation of texture. The second type of rotation results in a different projection of the trabecular structures on the radiograph: the texture is slightly different. In our experiments we rotated the femora over 15 degrees, perpendicular to the plane of the radiograph. Although due to repositioning variations, rotations in the plane will be present, the effect of the second type of rotation is made relatively stronger. Since we consider a rotation of 15 degrees as an upper limit for the rotational error in repositioning, the criteria for reproducibility, described in Table 2, were somewhat relaxed. As long as the variation due to rotation was under 50% of the subject variance, the reproducibility was graded as good: moderately independent of rotation. Since we are only interested in the effect of rotation on texture features which we consider reproducible under repositioning, a restricted set of features, based on the repositioning experiment, was selected. Comparing the results for the different regions, it is clear that in region III the texture varies the most due to rotation. Therefore, when rotational errors are to be expected, region III is not suitable for texture analysis. The texture in region IV is more susceptible to rotation then that in region I or II, but this is dependent on the texture features used. We recommend that the selection of the region of interest should be guided by a sensitivity analysis such as just described. In our study, the co-occurrence features are not much influenced by small texture differences due to rotation, as occurs in region I, II and IV. For the Fourier features, the dependence varies with the definition of the features used. The same accounts for the fractal features and the morphology gradient features.

Except for the local Fourier features, of which we expected a low reproducibility, it is virtually impossible to indicate beforehand which features will strongly react to repositioning or rotation. In the first place, the visual differences between the radiographs due to repositioning are very difficult to describe. Secondly, for most of the texture analysis methods we used in this study, the correlation between the features and visual aspects of the image is not well known.

Overall, the reproducibility of most features used, is rather high. Therefore, we think that texture analysis can be used for the analysis of the structure of clinical bone radiographs. Although our study was limited to the femur, we feel that we can extrapolate these results to other trabecular sites in the skeleton. However, we strongly advise, before putting texture features to the use of discriminating between different diseases or different stages of a disease, to check whether, in the experimental situation chosen, the reproducibility of the texture features used is high enough for the purpose defined.

5. Conclusions

We studied the reproducibility of several texture features in radiographs. We found, in general, that the reproducibility of texture features in radiographs is dependent on the location of the region of interest, and the degree of rotation. More specifically, we found that most of the co-occurrence features, and most of the fractal features are well reproducible. For the Fourier features and the morphology gradient features, a subset of features show a high reproducibility.

We would like to stress the point that before evaluating the ability of texture features in discriminating between different disease conditions, it is always necessary to assess the reproducibility of the features. Even more so, when the radiographs are taken at large time intervals, or in different clinical centers. Since the digitizing device has a profound influence on the texture content of the image, the digitizing conditions should be kept as constant as possible. Since texture features are sensitive to differences in exposure conditions, differences in screen-film material, and differences in digitizing conditions, all these conditions should be kept as constant as possible.

References

1. Lynch JA, Hawkes DJ, Buckland-Wright JC. Analysis of texture in macroradiographs of osteoarthritic knees using the fractal signature. Phys Med Biol 1991;36(6):709-722.

2. Ruttimann ES, Webber RL, Hazelrig JB. Fractal dimension from radiographs of peridental alveolar bone. Oral Surg Oral Med Oral Path 1992;74:98-110.
3. Benhamou CL, Lespessailles E, Jacquet G, et al. Fractal organization of trabecular bone images on calcaneus radiographs. J Bone Miner Res 1994;9(12):1909-1918.
4. Buckland-Wright JC, Lynch JA, Rymer J, Fogelman I. Fractal signature analysis of macroradiographs measures trabecular organization in lumbar vertebrae of postmenopausal women. Calcif Tissue Int 1994;54:106-112.
5. Caligiuri P, Giger ML, Favus MJ, Jia H, Doi K, Dixon LB. Computerized radiographic analysis of osteoporosis: preliminary evaluation. Radiology 1993;186:471-474.
6. Caldwell CB, Willett K, Cuncins AV, Hearn TC. Characterization of vertebral strength using digital radiographic analysis of bone structure. Med Phys 1995;22(5):611-615.
7. Singh M, Riggs BL, Beabout JW, Jowsey J. Femoral trabecular pattern index for evaluation of spinal osteoporosis. Mayo Clin Proc 1973;48:184-189.
8. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern 1973;3(6):610-621.
9. Bajcsy R. Computer description of textured surfaces. Proc 3rd Int Joint Conf on Artificial Intelligence 1973;572-579.
10. Katsuragawa S, Doi K, MacMahon H. Image feature analysis and computer-aided diagnosis in digital radiography: classification of normal and abnormal lungs with interstitial disease in chest images. Med Phys 1989;16(1):38-44.
11. Liu S-S, Jernigan ME. Texture analysis and discrimination in additive noise. CVGIP 1990;49:52-67.
12. Pentland AP. Fractal-based description of natural scenes. IEEE Trans Pattern Anal Machine Intell 1984;6:661-674.
13. Peleg S, Naor J, Hartley R, Avnir D. Multiple resolution texture analysis and classification. IEEE Trans Pattern Anal Machine Intell 1984;6:518-523.
14. Serra J. Image analysis and mathematical morphology (vol 1). London: Academic Press, 1982.
15. Veenland JF, Grashuis JL, Van der Meer F, Beckers ALD, Gelsema ES. Estimation of fractal dimension in radiographs. Med Phys 1996;(4):585-594.

Chapter 6

Unraveling the role of structure and density in determining vertebral bone strength

J.F. Veenland[1,2], T.M. Link[3], W. Konermann[3], N. Meier[3], J.L. Grashuis[1,2], E.S. Gelsema[1]

[1] Department of Medical Informatics
[2] Department of Radiology, Erasmus University Rotterdam, The Netherlands
[3] Department of Clinical Radiology, Münster, Germany

Abstract

The strength of bone is determined not only by bone density but also by structure. Therefore, quantification of the structure in radiographs by texture features may result in a better prediction of fracture risk. Since in radiographs density and structure are strongly correlated, the predictive power of texture features should be corrected for the influence of BMD to determine the additional information conveyed by these features. In this study, we evaluated the predictive power of various texture features based on the Gray-Level Dependence Method and the Morphological Gradient Method.

This study was performed on 67 vertebrae obtained from 20 male and 12 female human cadaver thoracolumbar spines. BMD and area of the vertebral body were determined from QCT images and texture features were derived from direct magnification (DIMA) radiographs. The fracture force, measured under conditions simulating the in vivo situation, was corrected with the area of the vertebra to yield the fracture stress (FS).

Results of the study indicate that BMD correlates significantly with FS r=0.82 (p<0.001, N=24) and r=0.94 (p<0.001, N=43) for female and male vertebrae, respectively. Correlation coefficients of the investigated texture features were as high as 0.80 (p<0.001) and 0.69 (p<0.001) for the female and male vertebrae, respectively. Multiple regression analysis showed that in female vertebrae the addition of one texture feature to BMD results in a better prediction of strength. The multiple correlation coefficient was 0.87 (p<0.001) in this case. In male vertebrae BMD was the best predictor of fracture stress.

These results suggest that texture features, as measured in magnification radiographs, can predict bone strength. Whereas in all cases BMD is the best single predictor of bone strength, for women texture features contain useful additional information.

Key words: *Vertebral strength, Magnification Radiographs, Texture, Quantitative Computed Tomography*

1. Introduction

The strength of bone is determined by bone density and structure. Whereas the density has been studied as a major determinant for years, the importance of structure has only been recognized more recently. This can be attributed to the fact that structure constitutes a feature that is, compared to density, difficult to quantify. Several sophisticated techniques have become available to estimate bone density; methods that measure structure in vivo, are being developed. Several studies[1-5] have shown that bone structure, more specifically trabecular architecture, can be quantified using texture analysis methods on radiographs.

The trabecular structure, as projected on a radiograph, has been studied using fractal dimensions[1,2,4,5] and features derived from the Fourier Spectrum[3]. We found that several fractal dimension estimation methods are influenced by the radiographic process and therefore should only be used under certain restrictions[6]. In the present study we have used texture features based on the Spatial Gray-Level Dependence Method (SGLDM) and on the Morphological Gradient Method. In the first method, gray-level differences over a short distance in a given direction are quantified. The texture features based on morphological gradients are sensitive to differences in direction, size, and shape of structures present in the image. The features, generated by either method, are true texture features: they are insensitive to first-order gray-level statistics such as the average optical density.

The contribution of trabecular bone to strength can be studied most directly in vertebrae, since these consist mainly of trabecular bone. In several studies, the strength of vertebrae has been related to an estimator of bone density, namely ash density[7], DPA[8,9], DXA[10-13] and QCT[9,11,13-19]. Also, histological parameters, quantifying the structure of trabecular bone, have been used in predicting the strength of vertebrae[20,21]. Although Caldwell et al.[22] showed that a structural feature, derived from the Sobel edge enhanced image, correlated significantly with compressive strength of the vertebrae, the value of this structural feature additional to bone density in predicting strength has not yet been studied. Since structure and density are two concepts that are closely intertwined physically and therefore strongly correlated, it is important to investigate whether structural features add any new information, apart from the density information which structural parameters inherently contain.

In this study we first investigated the correlation between different texture features, determined on direct magnification radiographs, and the strength of vertebrae. Second, the additional value of texture features in predicting bone

strength, independent of BMD, was evaluated for the male and female vertebrae separately.


2. Material and Methods

Thoracolumbar spines of 32 donors, 20 males and 12 females were obtained 1 or 2 days post mortem. Average age at death of the males and females was 60.2 and 56.6 years, respectively, ranging from 54 to 72 for the males and from 48 to 75 for the females. The spines were stripped of soft tissue, leaving discs, ligaments, and joint capsules intact. Donors with metastatic cancer or a history of trauma, as well as all vertebrae displaying defects on gross pathologic examination, were excluded. In this way 67 motion segments consisting of one vertebral body, the adjacent discs, and parts of the vertebrae were obtained from these spines. Of each segment, the midvertebral QCT and area was obtained, a magnification radiograph was taken, and the fracture force was determined.

QCT
All QCT examinations were performed using a CT scanner (Tomoscan LX, Philips) in an environment simulating the spine within the human body. Midvertebral scans with a slice thickness of 10 mm and 120 kVp were obtained. BMD was measured using a standardized ellipsoid region of interest (ROI). Additionally, the area of the vertebral body in the midvertebral scan was determined.

Magnification Radiography
Direct magnification (DIMA) radiographs of all specimens were obtained using the prototype of a microfocal X-ray tube (Microfox G10, Feinfocus Medical Systems, Garbsen, Germany). The images were taken with 6-fold magnification, a focal spot size of 60 mm and 50 kVp. Computed radiography was used as an imaging system employing the storage phosphor technique (FCR-7000 storage phosphor unit, Fuji, Tokyo). The matrix size of the cassette was 1760x2140 with an element size of 200 mm x 200 mm. In Fig. 1 the DIMA radiograph of a vertebra is shown.
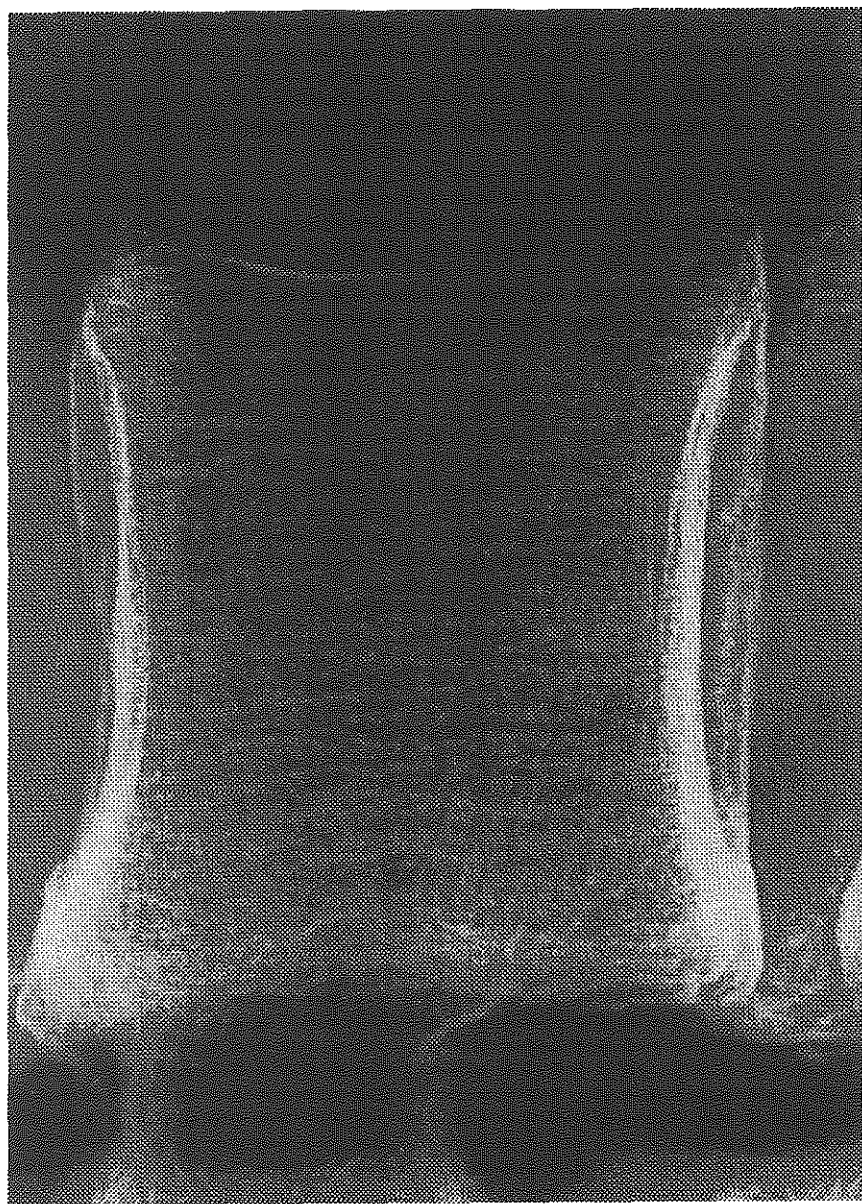
Fig 1. Direct magnification radiograph of a vertebra

Texture Features

In each DIMA radiograph, a region of interest of approximately 35 mm x 35 mm was selected centrally in the vertebra. This region was digitized using a CCD-camera, the resulting pixel size was 87.5 mm x 87.5 mm. In this study, we used two texture analysis methods: the Spatial Gray-Level Dependence Method[23] and the Morphological Gradient Method; both methods are described in Chapter 2.

Spatial Gray-Level Dependence Method:

In this method, the first step consists of a histogram equalization in which the number of gray-levels is reduced. Thereafter, the co-occurrence features are determined at small scales of 1 to 4 pixels. However, histogram equalization has one drawback: when a background trend is present, this trend is enhanced at the expense of the more subtle gray-level variations at a smaller, spatial, scale. Performing a background correction ensures that, instead of the background trend, the smaller gray-level variations are quantified. We performed a simple background correction by using a large uniform filter.

The histogram of each image was equalized to four gray-levels. Then, for the computation of the intermediate matrix, three distances d were selected: values of 1, 2, and 4 pixels were used. For the angle θ, values of $0°$ and $90°$ were taken. As the last step, the five features, as described in Chapter 2 were computed for each distance d and angle θ. An overview of the features used is given in Table 1.

Morphological Gradient Method:

This method was developed by ourselves especially for use in radiographs. The features are defined in Chapter 2, the different sizes of the structuring element used, are given in Table 1.

Table 1. Texture features generated by the two texture analysis methods

Method	Features
SGLDM	$ASM_{d,\theta}$, $CON_{d,\theta}$, $COR_{d,\theta}$, $DMO1_{d,\theta}$, $ENT_{d,\theta}$
	for $(d,\theta)=(1,0°), (2,0°), (4,0°), (1,90°), (2,90°), (4,90°)$
MGM	GR_5, GR_7, GR_9, GR_{11}
	$HSS_{7,3}$, $HSS_{9,5}$, $HSS_{11,7}$, $HSS_{13,9}$, $HSS_{15,11}$, $HSS_{17,13}$
	$HO_{5,1}$, $HO_{7,1}$, $HO_{7,3}$, $HO_{9,3}$, $HO_{11,3}$, $HO_{11,5}$, $HO_{13,3}$, $HO_{13,5}$

The features generated by both methods are independent of first-order gray-level statistics. Whereas the SGLD Method obtains its independence by the histogram equalization step, the features of the MG Method are inherently invariant under linear gray-scale transformations.

The reproducibility of the different texture features was determined on femora[24]. Only the features which showed a good reproducibility on femora were used in this study.

Mechanical Testing

Bone strength was determined as described by Brinckmann et al.[16]. The fracture force was obtained using an electro-mechanical materials testing machine in which the specimens were loaded. The specimens were mounted in a metal cup by means of a small amount of high viscosity bone cement underneath the endplate. The upper endplate was also covered with bone cement and a stainless steel plate was mounted on it. The surface of the plate was aligned parallel to the plane of the metal cup. The load was applied on the top of the mounted specimen by means of a stamp with a flat end. Compressive force on the specimens was increased gradually. Load and compressive deformation of the specimen were recorded on an xy-plotter. The curve exhibited an almost linear increase of deformation with force. Fracture occurred when the curve deviated from this linear elastic region and was accompanied by an audible cracking and extrusion of bone marrow through the ventral and dorsal venous pathways of the vertebral body of the motion segment. Fracture stress was calculated by dividing the fracture force by the midvertebral area, determined by CT. The midvertebral area was used, since the fracture load is distributed over this minimal cross-section.

Statistical Analysis

Pearson's product moment correlation coefficient was used to determine bivariate correlation coefficients. The correlation between two variables can be caused by the dependence of both variables on a third feature. In that case, the partial correlation coefficient can be used to compute the correlation between the two variables, independent of the third parameter. When r_{xy}, r_{xz}, and r_{yz} are the pair-wise correlation coefficients between the variables x, y, and z, the partial correlation coefficient between y and z corrected for x, is given by[25]

$$r_{yz.x} = \frac{r_{yz} - r_{xy}\, r_{xz}}{\sqrt{(1 - r_{xy}^2)(1 - r_{xz}^2)}} \tag{1}$$

115

In order to determine the correlation between a linear combination of two variables (e.g. x and y) and a third parameter (e.g. z), the multiple correlation coefficient ($R_{z.xy}$) was used. The multiple correlation coefficient is defined as

$$R_{z.xy} = \sqrt{\frac{r_{xz}^2 + r_{yz}^2 - 2\, r_{xy}\, r_{xz}\, r_{yz}}{1 - r_{xy}^2}} \qquad (2)$$

The relationship between the multiple correlation coefficient $R_{z.xy}$ and the partial correlation coefficient $r_{yz.x}$ is described in

$$R_{z.xy}^2 = r_{xz}^2 + r_{yz.x}^2 (1 - r_{xz}^2) \qquad (3)$$

This equation shows that the partial correlation coefficient determines whether combining the variable y with x will improve the prediction of parameter z.

In our study, we used the partial correlation coefficient for separating the contribution of BMD and structure features to fracture stress. Linear relations between the texture features, BMD, and fracture stress were assumed. The different analyses were performed for the whole set of vertebrae, and for both sexes separately.

3. Results

Table 2 shows the means, standard deviations and ranges of BMD, area and fracture force for the three vertebral categories. Since a motion segment also comprises the two adjacent vertebrae, at most one vertebra of each individual is present in each category. Whereas the BMD reaches the highest values in category T11-T12, area and fracture force increase in the caudal direction. Fracture force is divided by area, to correct for the influence of the size of the vertebra. The resulting parameter is called fracture stress.

The correlation between fracture stress and BMD for male and female vertebrae is depicted in Figs. 2(a) and 2(b), respectively. The correlation for the female vertebrae is somewhat lower than for the male vertebrae (r=0.82 and r=0.94, respectively). In Figs. 3(a) and 3(b) the correlation between fracture stress and a texture feature based on morphological gradients (GR$_9$) is illustrated. For female vertebrae, the correlation of the texture feature with fracture stress is almost as high

Table 2. BMD, area, and fracture force for the three vertebrae categories

	N	BMD			Area			Fracture Force		
		Mean	SD	Range	Mean	SD	Range	Mean	SD	Range
T11 T12	26	135.61	42.48	62.94-215.9	11.73	2.31	7.23-16.30	4.94	1.88	1.50-8.35
L1 L2	20	128.59	40.92	71.80-197.21	12.09	2.22	7.98-16.20	5.08	2.11	1.63-9.15
L3 L4	21	120.78	37.03	56.97-188.09	12.75	2.23	8.71-15.90	5.34	1.72	2.38-9.04

(r=0.79) as the correlation of BMD with fracture stress, whereas in male vertebrae the correlation is lower (r=0.51).

The correlation coefficients of several texture features with both fracture stress and BMD are listed in Table 3, 4, and 5 for all, male, and female vertebrae, respectively. In addition, the correlation of texture features with fracture stress, controlling for BMD is shown. For ease of survey each group of texture features is represented by one feature.

The texture features based on morphological gradients, with the exception of the orientation feature, showed a better correlation with fracture stress and BMD than the features based on the SGLD Method. The correlation with both fracture stress and BMD was higher for the female vertebrae than for the male vertebrae. Correction for the influence of BMD, using partial regression analysis, lowered the correlation coefficients substantially. Three texture features, all based on morphological gradients, were still significantly correlated with fracture stress in female vertebrae, whereas in male vertebrae all correlation coefficients became non-significant. On the other hand, the correlation between BMD and fracture stress can also be corrected for the influence of texture features. Using $HSS_{13,9}$ as a texture feature, in male vertebrae the correlation was only slightly lowered (r=0.89, p<0.001), whereas in female vertebrae this correction resulted in a substantially lower partial correlation coefficient of r=0.52 (p<0.05).

The result of multiple regression analysis can be calculated using eq. 3. In male vertebrae, combining BMD with texture features did not improve the prediction of fracture stress. In female vertebrae, however, the combination of BMD and one texture feature resulted in higher correlations (R=0.87, R=0.86, R=0.86 for GR_9, $HSS_{13,9}$, and $HO_{11,3}$ respectively) than for BMD alone (r=0.82).

117
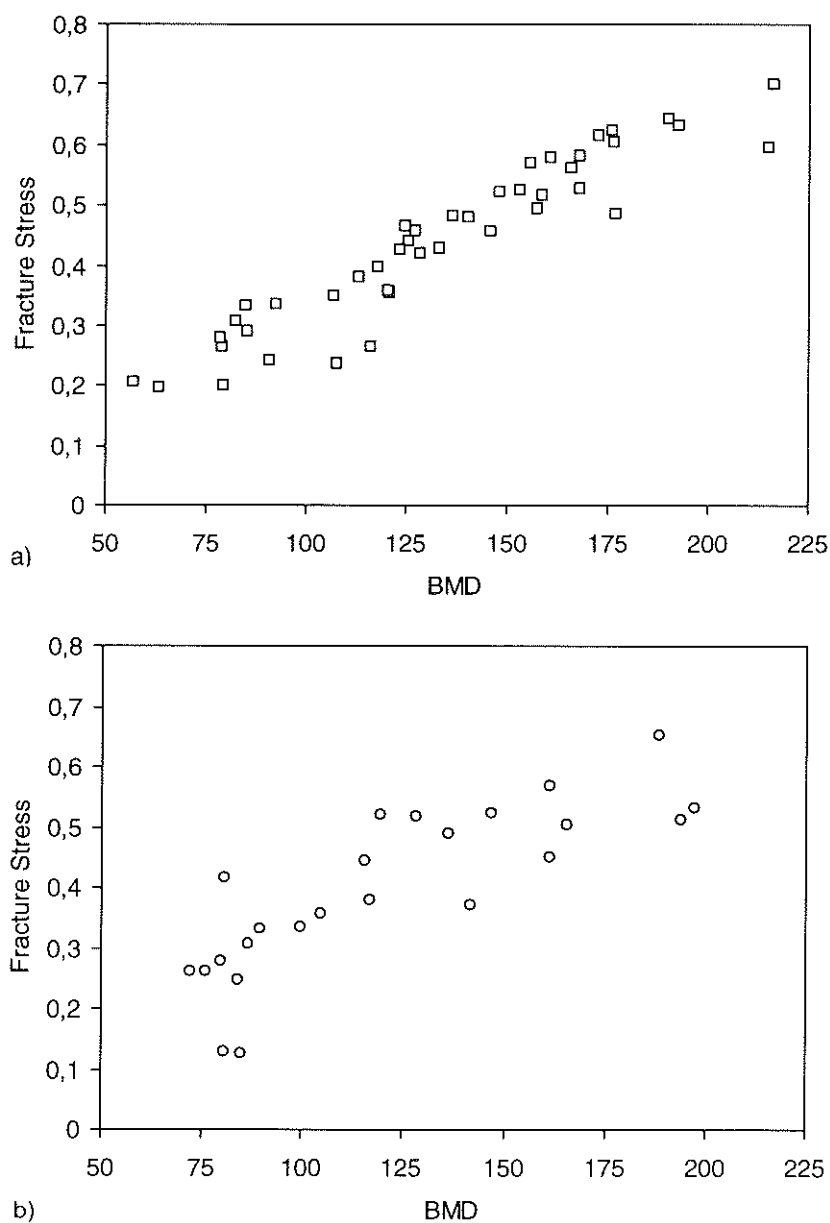
Fig 2. Fracture stress versus BMD for male (a) and female vertebrae (b).
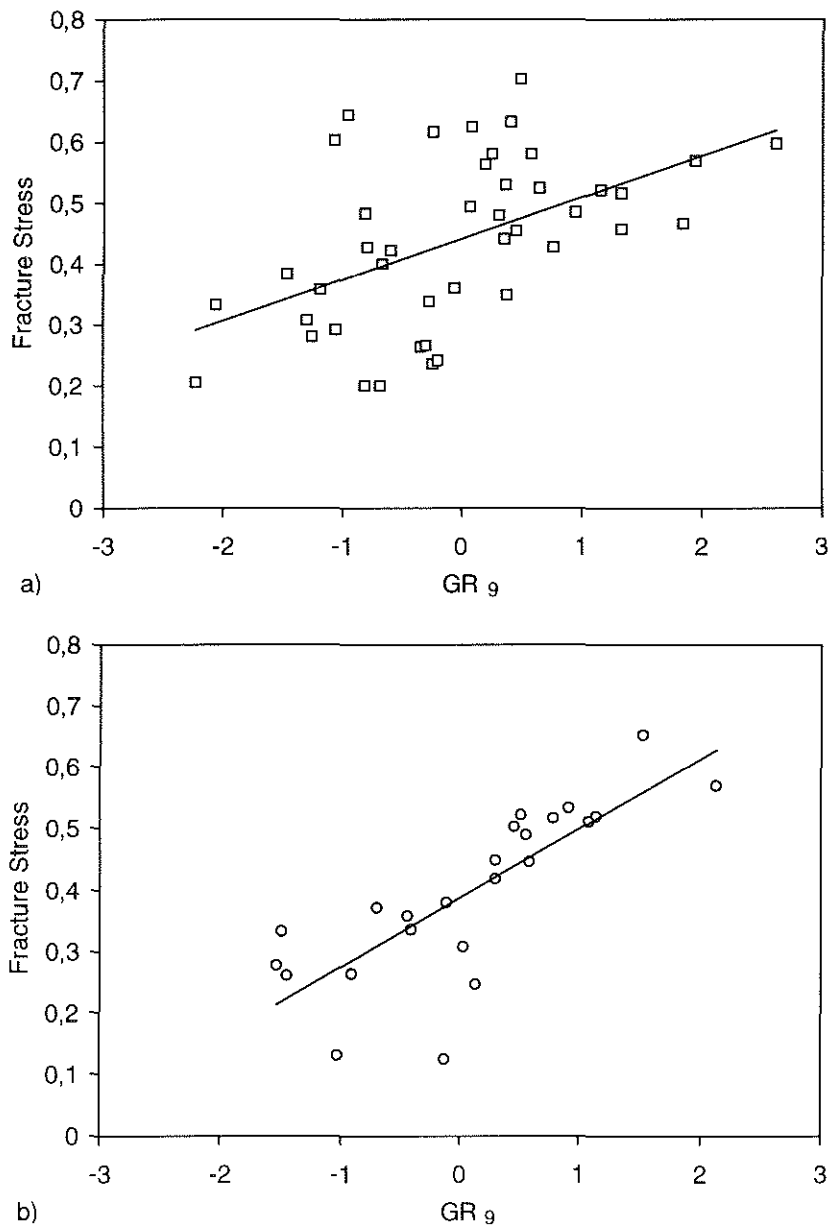
Fig 3. *Fracture stress versus the texture parameter GR_9, based on morphological gradients, for male (a) and female vertebrae (b).*

Table 3. *Correlation coefficient for texture parameters with fracture stress, BMD and fracture stress corrected for BMD for all vertebrae.*

	Fracture Stress	BMD	Fracture Stress ∣ BMD
$COR_{1,0°}$	0.466^a	0.440^a	0.179
$COR_{2,0°}$	0.455^a	0.419^a	0.197
$COR_{1,90°}$	0.314^b	0.340^b	0.019
$COR_{2,90°}$	0.141	0.151	0.012
GR_9	0.585^a	0.604^a	0.116
$HSS_{13,9}$	0.708^a	0.739^a	0.144
$HO_{11,3}$	-0.440^a	-0.368^b	-0.270^c

^{a} p<0.001; ^{b} p<0.01; ^{c}p<0.05.

Table 4. *Correlation coefficient for texture parameters with fracture stress, BMD and fracture stress corrected for BMD for male vertebrae.*

	Fracture Stress	BMD	Fracture Stress ∣ BMD
$COR_{1,0°}$	0.426^b	0.430^b	0.069
$COR_{2,0°}$	0.420^b	0.429^b	0.051
$COR_{1,90°}$	0.317^c	0.366^c	-0.089
$COR_{2,90°}$	0.181	0.241	-0.142
GR_9	0.508^a	0.574^a	-0.118
$HSS_{13,9}$	0.688^a	0.742^a	-0.049
$HO_{11,3}$	-0.303^c	-0.255	-0.195

^{a} p<0.001; ^{b} p<0.01; ^{c} p<0.05.

Table 5. *Correlation coefficient for texture parameters with fracture stress, BMD and fracture stress corrected for BMD for female vertebrae.*

	Fracture Stress	BMD	Fracture Stress I BMD
$COR_{1,0°}$	0.552^b	0.479^c	0.317
$COR_{2,0°}$	0.540^b	0.433^c	0.358
$COR_{1,90°}$	0.318	0.294	0.141
$COR_{2,90°}$	0.036	-0.073	0.168
GR_9	0.790^a	0.719^a	0.503^c
$HSS_{13,9}$	0.804^a	0.777^a	0.461^c
$HO_{11,3}$	-0.703^a	-0.590^b	-0.473^c

[a] $p<0.001$; [b] $p<0.01$; [c] $p<0.05$.

4. Discussion

Several studies have been performed relating density and strength of vertebrae. Due to the variety of parameters used, i.e., ash density, BMD determined by DPA, DXA, and QCT for density, and ultimate force, yield strength, and absorbed energy for strength, the results of these studies are difficult to compare.

The strength of vertebrae is determined by geometry, density, and structure. The contribution of these three factors should be evaluated separately. In Table 2. it can be seen that the fracture force increases with the area of the vertebrae in the caudal direction, whereas the BMD decreases. In this study, we assumed a proportional relation between area of the vertebra and strength, so as to eliminate the influence of geometry: the fracture force is divided by the area of the vertebra, resulting in the parameter fracture stress. Texture features are used to quantify the structure of trabecular bone, independent of first order gray-level statistics such as the average gray-level. The exact way in which the texture features are related to the underlying three-dimensional structure of the trabecular bone is still unclear. We chose to evaluate the vertebrae of men and women separately, since the female vertebrae are reported to contain relatively more trabecular bone than male vertebrae[26].

Since bone mineral density and trabecular structure are two concepts that are closely intertwined, no structure without density and no density without structure,

121

we used partial correlation techniques to investigate the influence of both factors separately. In partial correlation, the effects of other variables can be controlled by removing the linear relationship of these variables with the variables studied. In this way, in the present study, the relationship between structure and strength was studied, controlling for the effect of density. Also, the relation between density and strength was quantified, independent of structure. The partial correlation coefficient will depend on the texture feature used.

We found that the correlation between texture features and fracture stress was lower for male than for female vertebrae. Both in male and female vertebrae this correlation is partly based on BMD (Tables 3, 4, and 5). After removing the influence of BMD, only in female vertebrae the texture features based on morphological gradients are still significantly correlated. Therefore, it can be stated that the features based on the MG Method, in contrast to those based on the SGLD Method, are able to convey real structure information independent of BMD. Correction for the influence of the texture feature on the correlation between BMD and fracture stress, changed the correlation coefficient for male vertebrae only slightly, whereas the correlation coefficient for female vertebrae was substantially lowered. This means that the BMD, in female vertebrae, is partly determined by the trabecular structure.

Multiple regression analysis showed that in male vertebrae BMD alone is the best predictor of fracture stress; texture features do not add new information. In female vertebrae, however, a combination of one texture feature and BMD provides a better prediction of fracture stress than BMD alone.

These observed differences between male and female vertebrae may be explained by the different ratios of trabecular and cortical bone in the vertebrae of both sexes[26]: the female vertebrae contain relatively more trabecular bone. Therefore, it can be assumed that in female vertebrae the contribution of the trabecular bone to strength is more important than in male vertebrae. Going one step further, we conjecture that in female vertebrae the architecture of the trabecular bone is a more critical factor in determining the strength of the vertebra than in male vertebrae.

Based on these results, we can conclude that, at least in female vertebrae, trabecular structure as measured by texture features, is an important determinant of vertebral bone strength. Moreover, we have shown that part of the structure information is already included in the BMD measurement. Therefore, structure information has additive importance to BMD in predicting strength, but because of the interrelation of structure and density, this additive value is limited.

Since the sample size in this study is fairly small, the results will have to be validated on a larger test set.

References

1. Lynch JA, Hawkes DJ, Buckland-Wright JC. Analysis of texture in macroradiographs of osteoarthritic knees using the fractal signature. Phys Med Biol 1991;36(6):709-722.
2. Ruttimann ES, Webber RL, Hazelrig JB. Fractal dimension from radiographs of periodontal alveolar bone. Oral Surg Oral Med Oral Path 1992;74:98-110.
3. Caligiuri P, Giger ML, Favus MJ, Jia H, Doi K, Dixon LB. Computerized radiographic analysis of osteoporosis: preliminary evaluation. Radiology 1993;186:471-474.
4. Benhamou CL, Lespessailles E, Jacquet G, et al. Fractal organization of trabecular bone images on calcaneus radiographs. J Bone Miner Res 1994;9(12):1909-1918.
5. Buckland-Wright JC, Lynch JA, Rymer J, Fogelman I. Fractal signature analysis of macroradiographs measures trabecular organization in lumbar vertebrae of postmenopausal women. Calcif Tissue Int 1994;54:106-112.
6. Veenland JF, Grashuis JL, Van der Meer F, Beckers ALD, Gelsema ES. Estimation of fractal dimension in radiographs. Med Phys 1996;(4):585-594.
7. Mosekilde L, Mosekilde L, Danielsen CC. Biomechanical competence of vertebral bone in relation to ash density and age in normal individuals. Bone 1987;8:79-85.
8. Hansson T, Roos B, Nachemson A. The bone mineral content and ultimate compressive strength of lumbar vertebrae. Spine 1980;5(1):46-55.
9. Eriksson SAV, Isberg BO, Lindgren JU. Prediction of vertebral strength by dual photon absorptiometry and quantitative computed tomography. Calcif Tissue Int 1989;44:243-250.
10. Myers BS, Arbogast KB, Lobaugh B, Harper KD, Richardson WJ, Drezner MK. Improved assessment of lumbar vertebral body strength using supine lateral dual-energy X-ray absorptiometry. J Bone Miner Res 1994;9(5):687-693.
11. Edmondston SJ, Singer KP, Day RE, Breidahl PD, Price RI. In-vitro relationships between vertebral body density, size, and compressive strength in the elderly thoracolumbar spine. Clin Biomech 1994;9:180-186.
12. Moro M, Hecker AT, Bouxsein ML, Myers ER. Failure load of thoracic vertebrae correlates with lumbar bone mineral density measured by DXA. Calcif Tissue Int 1995;56:206-209.
13. Singer K, Edmondston S, Day R, Breidahl P, Price R. Prediction of thoracic and lumbar vertebral body compressive strength: correlations with bone mineral density and vertebral region. Bone 1995;17(2):167-174.

14. McBroom RJ, Hayes WC, Edwards WT, Goldberg RP, White III AA. Prediction of vertebral body compressive fracture using quantitative computed tomography. J Bone Joint Surg Am 1985;67(8):1206-1214.
15. Lang SM, Moyle DD, Clemson D, et al. Correlation of mechanical properties of vertebral trabecular bone with equivalent mineral density as measured by computed tomography. J Bone Joint Surg Am 1988;70-A(10):1531-1538.
16. Brinckmann P, Biggemann M, Hilweg D. Prediction of the compressive strength of human lumbar vertebrae. Clin Biomech 1989;2:S1-S27.
17. Mosekilde L, Bentzen SM, Ørtoft G, Jørgensen J. The predictive value of quantitative computed tomography for vertebral body compressive strength and ash density. Bone 1989;10:465-470.
18. Biggemann M, Hilweg D, Seidel S, Horst M, Brinckmann P. Risk of vertebral insufficiency fractures in relation to compressive strength predicted by quantitative computed tomography. Eur J Radiol 1991;13(1):6-10.
19. Hayes WC, Piazza SJ, Zysset PK. Biomechanics of fracture risk prediction of the hip and spine by quantitative computed tomography. Radiol Clin North Am 1991;29(1):1-18.
20. Mosekilde L, Mosekilde L. Iliac crest trabecular bone volume as predictor for vertebral compressive strength, ash density and trabecular bone volume in normal individuals. Bone 1988;9:195-199.
21. Dempster DW, Ferguson-Pell MW, Mellish RWE, et al. Relationships between bone structure in the iliac crest and bone structure and strength in the lumbar spine. Osteoporosis Int 1993;3(2):90-96.
22. Caldwell CB, Willett K, Cuncins AV, Hearn TC. Characterization of vertebral strength using digital radiographic analysis of bone structure. Med Phys 1995;22(5):611-615.
23. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern 1973;3(6):610-621.
24. Veenland JF, Grashuis JL, Link TM, Gelsema ES. Reproducibility of texture features in femur radiographs. Proceedings of international workshop "Trabecular bone texture analyis on radiographs", Orleans, December 18-20, 1996.
25. Sachs L. Angewandte statistik. Berlin, Heidelberg: Springer-Verlag, 1984.
26. Nottestad SY, Baumel JJ, Kimmel DB, Recker RR, Heaney RP. The proportion of trabecular bone in human vertebrae. J Bone Miner Res 1987;2(3):221-229.

Part III:

Clinical Studies

Chapter 7

Texture differences between osteoporotic and non-osteoporotic trabecular bone independent of BMD

Abstract

Osteoporotic bone is characterized by a low bone mineral density and architectural changes. Whereas BMD is widely regarded as the main characteristic of osteoporosis, more recently research has been aimed at quantifying bone structure. Texture features can be used for the quantification of the trabecular bone structure in radiographs. However, structure and bone mineral density are two concepts which are closely intertwined. Consequently, texture features are inherently correlated with BMD. Since the goal of using texture features is not to replace density measurements, but to add structure information, the value of texture features in discriminating between osteoporotic and non-osteoporotic bone, should be evaluated independent of BMD.

In this study, we measured BMD in hand-radiographs of 40 osteoporotic and 40 non-osteoporotic females, using a technique based on radiogrammetry. Using these density values, 15 osteoporotic and 15 non-osteoporotic females were matched. Texture features were determined in the same hand-radiographs applying two texture analysis methods: the Spatial Gray-Level Dependence Method (SGLDM) and the Morphology Gradient Method (MGM). A Wilcoxon matched pairs signed rank sum test was performed to test whether the texture features differed for the matched pairs.

Results of the study indicate, that most of the MGM features and a number of SGLDM features are significantly different within the matched pairs. On average, the significance level of the MGM features was higher.

These results suggest that applying texture analysis to radiographs, differences can be measured between the osteoporotic and non-osteoporotic trabecular bone pattern independent of BMD.

1. Introduction

Osteoporotic bone is characterized by a low bone density and structural changes in the bone tissue. Several sophisticated techniques, such as QCT and DXA, have been developed to determine the bone density. Both methods have become common practice to estimate the fracture risk. However, a large overlap can be found in bone mineral densities of women with and without fractures[1-3]. Differences in bone structure might be one of the explanations. Therefore, new research efforts have been directed at the quantification of the structure of bone. The architectural changes in the trabecular bone have been described thoroughly using histomorphometric techniques[2,4-6]. A less invasive technique for studying trabecular structure is offered by radiography[7]: the 3-dimensional structure is projected as a two-dimensional texture in the radiograph. Since visual grading of trabecular textures in radiographs has proven to be subject to large inter- and intra-observer variability, a more objective, computerized, approach is preferred. Several authors have applied different texture analysis methods for the quantification of trabecular texture[8-16].

One of the problems encountered when trying to select or define a texture feature which is able to discriminate between osteoporotic and normal bone is the interrelationship between structure and density. In trabecular bone there is no structure without density and no density without structure. Despite the fact that the exact relation between the trabecular bone structure and the texture features derived from the radiograph is not known, a strong correlation between texture features and density can be observed, especially when measured at the same skeletal site[16]. The goal of using texture features is not to replace density measurements, but to quantify structural information. Therefore, texture features should be evaluated on their ability to discriminate between osteoporotic and non osteoporotic trabecular textures, independent of BMD.

In this study we investigated the textural differences between the radiographic trabecular pattern of osteoporotic females and controls, which were matched for density. The BMD was measured at the phalanx at which also the texture analysis was performed.

2. Material and Methods

Material

Standardized posterior-anterior (PA) and lateral radiographs (65 kV, 15mAs, 3M GT X-ray film and Trimax α2 intensifying screen) were made of the right hand of 40 healthy and 40 osteoporotic females. The diagnosis osteoporosis was based on the presence of vertebral fractures visible on lateral spine radiographs. Using the hand radiographs, the bone mineral density was determined on the phalanx media of the right index finger using the Quantitative MicroDensitometry method as described by Trouerbach et al.[17]. This method consists of a number of steps. In the first place, the intensity-values in the image are calibrated with the aid of a, simultaneously radiographed, aluminium wedge. Then the cross-section of, a section of, the bone is estimated by combining measurements of the PA and the lateral radiograph of the right index finger. Thereafter, the average, calibrated, intensity in the section mentioned is measured in the two radiographs. This bone mass measurement is then corrected with the cross-section of the bone, resulting in a bone mineral density measurement. Based on these density values, the osteoporotic and healthy females were matched. This matching procedure resulted in 15 pairs of handradiographs. For each pair, the squared difference of BMD was smaller than 0.5% of the variance of BMD as measured in the healthy females.

Texture Features

In these 30 PA radiographs, the texture features were determined on a region of interest of approximately 4.2x4.2 mm in the proximal part of the phalanx proximalis of the right index finger. In the region chosen, the trabecular structure appeared visually as homogenous. The regions were digitized (140 x 140 pixels) using an 8 bits CCD camera. In this study, we used two texture analysis methods: the Spatial Gray-Level Dependence Method[18] and the Morphological Gradient Method; both methods are described in Chapter 2.

Spatial Gray-Level Dependence Method

In this method, the first step consists of a histogram equalization in which the number of gray-levels is reduced. Thereafter, the co-occurrence features are deter-mined at small scales of 1 to 4 pixels. However, histogram equalization has one drawback: when a background trend is present, this trend is enhanced at the expense of the more subtle gray-level variations at a smaller, spatial, scale. Performing a background correction ensures that, instead of the background trend,

the smaller gray-level variations are quantified. We performed a simple background correction by using a large uniform filter (61x61 pixels).

The histogram of each image was equalized to four gray-levels. Then, for the computation of the intermediate matrix, three distances d were selected: values of 1, 2, and 4 pixels were used. For the angle θ, values of $0°$ and $90°$ were taken; the direction $0°$ was chosen in line with the scan direction of the CCD-camera. As the last step, the five features, as described in Chapter 2, were computed for each combination of distance d and angle θ. An overview of the features used is given in Table 1. These features will be referred to as the co-occurrence features.

Morphological Gradient Method
This method was developed by ourselves, especially for use in radiographs. The features are based on combinations of the basic operations of Mathematical Morphology[19]. Two morphological gradients are used: the G-gradient and the H-gradient. A complete description is given in Chapter 2. The features, which will be referred to as the morphological features, are listed in Table 1. The features *HR* and *GR* are sensitive to the roundness of structures and edges, respectively. The features *HSS* and *HSR* assess the size of structures and are derived with square and round structuring elements, respectively. *GSS* and *GSR* perform the same function for edges. The orientation of structures and edges present in the image are quantified by *HO* and *GO*.

Texture is constituted by the coherence, not the absolute values, of gray-levels. Therefore, texture features should be independent of first order gray-level statistics such as the average gray-level and the variance of the gray-levels. An ideal texture analysis method is invariant under strict monotonic gray-level transformations[18], however very few methods, e.g. the SGLD Method, meet this strict requirement. A less strict requirement can be used when the shape of the gray-level distributions is similar, namely invariance under linear transformations of the gray-levels. The features of the MG Method are by definition invariant under linear gray-scale transformations. The features of the SGLD Method obtain their invariance under monotonic gray-scale transformations by the histogram equalization step.

Statistical analysis
For each feature, a Wilcoxon matched pairs signed rank sum test was performed to test whether the texture features differed within the matched pairs. Using this test, no assumptions are made about the distribution of the texture features.

__Table 1.__ Texture features generated by the two texture analysis methods.

Method	Features
SGLDM	$ASM_{d,\theta}$, $CON_{d,\theta}$, $COR_{d,\theta}$, $DMO1_{d,\theta}$, $ENT_{d,\theta}$ for $(d,\theta)=(1,0°),(2,0°),(4,0°),(1,90°),(2,90°),(4,90°)$
MGM	GR_5, GR_7, GR_9, GR_{11}
	$GSS_{7,3}$, $GSS_{9,3}$, $GSS_{11,5}$
	$GSR_{7,3}$, $GSR_{9,3}$, $GSR_{11,5}$
	$GO_{7,1}$, $GO_{9,3}$, $GO_{11,3}$, $GO_{11,5}$, $GO_{13,5}$
	HR_5, HR_7, HR_9, HR_{11}
	$HSS_{7,3}$, $HSS_{9,3}$, $HSS_{11,5}$
	$HSR_{7,3}$, $HSR_{9,3}$, $HSR_{11,5}$
	$HO_{7,1}$, $HO_{9,3}$, $HO_{11,3}$, $HO_{11,5}$, $HO_{13,5}$

3. Results

In Table 2, the Z-scores as computed by the Wilcoxon matched pairs test are listed, together with their significances. For the co-occurrence features, it can be seen that the z-scores of the features computed perpendicular to the scanning direction of the CCD-camera (direction 90°) show a strong decrease with the distance d at which the features are determined. This decrease cannot be seen for the features computed in line with the scanning direction of the CCD-camera (direction 0°). Furthermore, for both directions and all distances, the features Contrast (*CON*) and First Diagonal Moment (*DMO1*) show higher Z-scores than the other features.

For the morphological gradient features, Table 3 shows that for most of the features used in this study, a significant difference can be found between the individuals forming a matched pair. Only the *HSR*-features are, apart from one, not significantly different between the matched pairs. Furthermore, it can be seen that the orientation features (*GO* and *HO*) show somewhat lower Z-scores than the other features used in this study.

Table 2. *Z-values and their significance as computed by the Wilcoxon matched pairs rank sum test of the co-occurrence features.*

Feature	Z	P	Feature	Z	P
$ASM_{1,0°}$	1.93	N.S.	$ASM_{1,90°}$	1.36	N.S.
$CON_{1,0°}$	2.07	<0.05	$CON_{1,90°}$	2.15	<0.05
$COR_{1,0°}$	1.68	N.S.	$COR_{1,90°}$	1.90	N.S.
$DMO1_{1,0°}$	2.04	<0.05	$DMO1_{1,90°}$	2.16	<0.05
$ENT_{1,0°}$	1.93	N.S.	$ENT_{1,90°}$	2.04	<0.05
$ASM_{2,0°}$	1.48	N.S.	$ASM_{2,90°}$	0.82	N.S.
$CON_{2,0°}$	2.10	<0.05	$CON_{2,90°}$	1.65	N.S.
$COR_{2,0°}$	1.79	N.S.	$COR_{2,90°}$	1.29	N.S.
$DMO1_{2,0°}$	2.04	<0.05	$DMO1_{2,90°}$	1.68	N.S.
$ENT_{2,0°}$	1.82	N.S.	$ENT_{2,90°}$	1.39	N.S.
$ASM_{4,0°}$	1.54	N.S.	$ASM_{4,90°}$	0.45	N.S.
$CON_{4,0°}$	2.04	<0.05	$CON_{4,90°}$	0.51	N.S.
$COR_{4,0°}$	1.87	N.S.	$COR_{4,90°}$	0.20	N.S.
$DMO1_{4,0°}$	1.96	N.S.	$DMO1_{4,90°}$	0.34	N.S.
$ENT_{4,0°}$	1.82	N.S.	$ENT_{4,90°}$	0.14	N.S.

Table 3. *Z-values and their significance as computed by the Wilcoxon matched pairs rank sum test of the morphological gradient features.*

Feature	Z	P	Feature	Z	P
GR_5	2.90	<0.01	HR_5	1.99	<0.05
GR_7	2.61	<0.01	HR_7	3.01	<0.01
GR_9	2.39	<0.05	HR_9	3.01	<0.01
GR_{11}	1.93	N.S.	HR_{11}	2.56	<0.05
$GSS_{7,3}$	3.07	<0.01	$HSS_{7,3}$	2.73	<0.01
$GSS_{9,3}$	3.07	<0.01	$HSS_{9,3}$	3.07	<0.01
$GSS_{11,3}$	2.89	<0.01	$HSS_{11,3}$	3.24	<0.01
$GSS_{13,3}$	2.78	<0.01	$HSS_{13,3}$	3.12	<0.01
$GSR_{7,3}$	3.07	<0.01	$HSR_{7,3}$	2.04	<0.05
$GSR_{9,3}$	3.07	<0.01	$HSR_{9,3}$	1.93	N.S.
$GSR_{11,3}$	2.95	<0.01	$HSR_{11,3}$	1.70	N.S.
$GSR_{17,3}$	2.78	<0.01	$HSR_{13,3}$	1.65	N.S.
$GO_{7,1}$	2.22	<0.05	$HO_{13,1}$	2.27	<0.05
$GO_{9,1}$	2.10	<0.05	$HO_{13,3}$	2.33	<0.05
$GO_{9,3}$	2.33	<0.05	$HO_{13,5}$	2.33	<0.05
$GO_{11,3}$	2.27	<0.05	$HO_{17,1}$	2.44	<0.05
$GO_{11,5}$	2.22	<0.05	$HO_{17,3}$	2.56	<0.05

4. Discussion

Osteoporosis is characterized by a change in bone architecture as well as by a reduced bone mineral density (BMD), resulting in an increased fracture risk. At present, BMD is used as the main predictor of fracture risk. However, it has been shown that the architectural changes in the trabecular bone are reflected in radiographs[7], and that features expressing texture can be used to extract structure information from radiographs[8-16]. One of the problems encountered, when trying to

quantify trabecular structure from radiographs, is that BMD and trabecular structure are closely interrelated: in trabecular bone there is no structure without density and no density without structure. Despite the fact that the exact relation between the trabecular bone structure and the texture features derived from the radiograph is not known, a strong correlation between texture features and density can be observed, especially when measured at the same skeletal site[16]. Some authors show explicitly that texture features can discriminate between bone with high and bone with low mineral density[13,20] or correlate significantly with BMD[15]. However, the purpose of applying texture analysis is not to replace the bone mineral density measurement: the bone mineral density is regarded widely as an important determinant of fracture risk. The role of texture features is an auxiliary one. Texture features reveal structural information which, when combined with bone mineral density, can contribute to a better prediction of fracture risk.

In a former study[16] we investigated the correlation of texture features with the strength of vertebral bone. For 67 vertebrae, the fracture force, and the BMD, using QCT, were measured. The texture features of the Spatial Gray-Level Dependence Method and the Morphological Gradient Method were determined in a region located centrally in the vertebrae. We found that both the Spatial Gray-Level Dependence Method and the Morphological Gradient Method yielded a number of features correlating significantly with the strength of vertebral bone. However, when the correlation was corrected for BMD, only the morphologic gradient features showed a significant correlation with strength.

The goal of this chapter was to investigate the usefulness of these two sets of texture features in discriminating osteoporotic from non osteoporotic trabecular texture, independent of BMD. For that purpose osteoporotic and non-osteoporotic females were matched for BMD. The BMD was measured in the same phalanx in which also the trabecular structure was quantified.

Two texture analysis methods were selected for this study: the Spatial Gray-Level Dependence Method and the Morphology Gradient Method. The features of both methods are independent of the average gray-level and the variance of the gray-levels. In this way, the influence of differences in optical density due to, e.g., variations in exposure time on texture features are minimized. Furthermore, we showed[21] that both methods are, as compared to other texture analysis methods, less influenced by image blur, which is caused by the intensifying screen.

We found that a number of co-occurrence features, and almost all morphologic gradient features were significantly different between the individuals forming a density-matched pair. On average, the level of significance of the morphology gradient features was higher than that of the co-occurrence features. The features

135

differed in their Z-value, but since the relation between the texture features and the visible texture is still unclear, it cannot be readily explained why some features perform better than others. Since the CCD-camera has a modulation transfer function in the scanning direction which differs from the one in the direction perpendicular, different behaviour of the co-occurrence features for those two directions can be expected. The signal in the scanning direction is low-pass filtered, the higher frequencies are eliminated. This can explain the consistent behaviour of the co-occurrence features over different distances in the scanning direction of the CCD camera.

In this study, we have focussed on quantifying texture differences in hand radiographs. When using texture analysis methods on radiographs for the quantification of trabecular bone, the number of suitable skeletal sites is limited. Clinical radiographs of the spine and femur are not very suitable for texture analysis due to low spatial resolution, the (differences in) superimpositions, differences in object-focus distance, and the low repositioning reproducibility. It has been shown that radiographs of wrists and hands are suitable for texture analysis22. There are no superimpositions, the object-focus distance is fairly constant, repositioning reproducibility is high, and, furthermore, there is the possibility to use a slightly higher dose and a sharper screen-film system for a higher spatial resolution.

We can conclude that texture differences can be measured between osteoporotic and non-osteoporotic trabecular bone, independent of BMD. Since the sample size in this study is very limited, the results will have to be validated on a larger set.

References

1. Ross PD, Heilbrun LK, Wasnich RD, Davis JW, Vogel JM. Methodologic issues in evaluating risk factors for osteoporotic fractures. J Bone Miner Res 1989;4(5):649-656.
2. Kimmel DB, Recker RR, Gallagher JC, Vaswani AS, Aloia JF. A comparison of iliac bone histomorphometric data in post-menopausal osteoporotic and normal subjects. Bone and Mineral 1990;11:217-235.
3. Ott SM. When bone mass fails to predict bone failure. Calcif Tissue Int 1993;53:S7-13.
4. Parfitt AM, Mathews CHE, Villaneuva AR, Kleerekoper M. Relationships between surface, volume, and thickness of iliac trabecular bone in aging and in osteoporosis. J Clin Invest 1983;72:1396-1409.

5. Mosekilde L. Age-related changes in vertebral trabecular bone architecture-assessed by a new method. Bone 1988;9:247-250.
6. Snyder BD, Edwards WT, Hayes WC. Role of trabecular morphology in the etiology of age-related vertebral fractures. Calcif Tissue Int 1993;53:S14-22.
7. Resnick D, Niwayama G. Diagnosis of bone and joint disorders. Saunders Company, 1988.
8. Lynch JA, Hawkes DJ, Buckland-Wright JC. Analysis of texture in macroradiographs of osteoarthritic knees using the fractal signature. Phys Med Biol 1991;36(6):709-722.
9. Geraets WGM, van der Stelt PF, Netelenbos CJ, Elders PJM. A new method for automatic recognition of the radiographic trabecular pattern. J Bone Miner Res 1990;5(3):227-233.
10. Ruttimann ES, Webber RL, Hazelrig JB. Fractal dimension from radiographs of peridental alveolar bone. Oral Surg Oral Med Oral Path 1992;74:98-110.
11. Caligiuri P, Giger ML, Favus MJ, Jia H, Doi K, Dixon LB. Computerized radiographic analysis of osteoporosis: preliminary evaluation. Radiology 1993;186:471-474.
12. Benhamou CL, Lespessailles E, Jacquet G, et al. Fractal organization of trabecular bone images on calcaneus radiographs. J Bone Miner Res 1994;9(12):1909-1918.
13. Buckland-Wright JC, Lynch JA, Rymer J, Fogelman I. Fractal signature analysis of macroradiographs measures trabecular organization in lumbar vertebrae of postmenopausal women. Calcif Tissue Int 1994;54:106-112.
14. Caldwell CB, Willett K, Cuncins AV, Hearn TC. Characterization of vertebral strength using digital radiographic analysis of bone structure. Med Phys 1995;22(5):611-615.
15. Lee RL, Dacre JE, James MF. Image processing assessment of femoral osteopenia. J of Digit Imag 1997;10(3 S 1):218-221.
16. Veenland JF, Link TM, Konermann W, Meier N, Grashuis JL, Gelsema ES. Unraveling the role of structure and density in determining vertebral bone strength. Calcif Tissue Int 1997;61(6):474-479.
17. Trouerbach WT, Birkenhager JC, Collette BJA, Drogendijk AC, Schmitz PIM, Zwamborn AW. A study on the phalanx bone mineral content in 273 normal pre- and post-menopausal females (transverse study of age-dependent bone loss). Bone and Mineral 1987;3:53-62.
18. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern 1973;3(6):610-621.
19. Serra J. Image analysis and mathematical morphology (vol 1). London: Academic Press, 1982.
20. Southard KA, Southard TE. Comparison of digitized radiographic alveolar features between 20- and 70-year-old women. Oral Surg Oral Med Oral Path 1992;74:111-117.

21. Veenland JF, Grashuis JL, Gelsema ES. Texture analysis in radiographs: the influence of MTF and noise on the discriminative ability of texture features. Med Phys 1998;25(6):922-936.
22. Geraets WGM, van der Stelt PF, Elders PJM. The radiographic trabecular bone pattern during menopause. Bone 1993;14:859-864.

Chapter 8

Suitability of texture features as indicators of vertebral bone strength

Abstract

The strength of bone is not only determined by bone mineral density but also by the structure. In osteoporosis, the loss of bone mass affects both: BMD, and the structure. Whereas BMD is widely regarded as the main characteristic of osteoporosis, more recently research has been aimed at quantifying bone structure. It has been shown that texture features can be used for the quantification of the trabecular bone structure in radiographs.

In this study, we used the presence of vertebral deformities as an indicator of decreased strength. Since clinical spine radiographs are not suitable for the determination of texture features, hand radiographs were used for this purpose. We applied two texture analysis methods: the Spatial Gray-Level Dependence Method (SGLDM) and the Morphology Gradient Method (MGM). We investigated, using the student's t-test and logistic regression, whether texture features, measured on hand radiographs, can indicate the strength of vertebral bone, independent of bone mineral density and age.

Results of the study indicate, that the majority of the MGM features show a significant difference between the group with vertebral deformities and the control group. None of the SGLDM features were significantly different between the two groups. Since these differences in texture could be related to differences in age or BMD, a multivariate logistic regression adjusting for age and BMD was performed. Most of the MGM features remained significant after correction for age and BMD.

In this study, we showed that morphological texture features, measured in hand radiographs, are significantly related, independent of age and BMD, to the presence of one or more deformities in the spinal column.

Key words: *Vertebral deformities, Texture, Osteoporosis*

1. Introduction

Osteoporosis is a condition characterized by a loss of bone mass, leading to an increased bone fragility and susceptibility to fracture. This loss of bone mass does not only affect the bone mineral density, but also the structure of bone. Whereas the density has been studied as a major characteristic of osteoporosis for years, the importance of structure has only more recently been recognized. This can be attributed to the fact that structure constitutes an attribute which is, as compared to density, more difficult to quantify. Several sophisticated techniques have become available to estimate bone density; methods which measure structure in vivo, are being developed. Several studies[1-10] have shown that bone structure, more specifically trabecular architecture, can be quantified using texture analysis methods on radiographs.

Since osteoporosis is a systemic disease, axial skeletal sites can be used to predict fracture risk at appendicular sites, and the other way round. However, for bone mineral density it has been shown that the predictive power is strongest when measured at the site at risk. When using methods of texture analysis on radiographs for the quantification of trabecular bone, the number of suitable skeletal sites is limited. Clinical radiographs of the spine and femur are not very suitable for texture analysis due to low spatial resolution, the (differences in) overprojections, differences in object-focus distance, and the low repositioning reproducibility. On the other hand, the radiographs of wrists and hands have shown to be very suitable for texture analysis[11]. There are no overprojections, object-focus distance is fairly constant, repositioning reproducibility is high, and, furthermore, in such appendicular sites it is possible to use a slightly higher dose and a sharper screen-film system for a higher spatial resolution than is possible in more axial sites.

In the present study we used two texture analysis methods on hand radiographs: the Spatial Gray-Level Dependence Method (SGLDM) and the Morphological Gradient Method. In the first method, gray-level differences over a short distance in a given direction are quantified. The texture features based on morphological gradients are sensitive to differences in direction, size and shape of the structures present in the image. The features, generated by either method, are true texture features: they are independent of first-order gray-level statistics such as the average optical density.

In a previous study[10], in which cadaver vertebrae were crushed, we found that several texture features, determined on radiographs of the uncrushed vertebrae, showed a significant correlation with the strength of bone, independent of BMD. However, in a clinical study there is no way to directly measure the strength of

vertebrae; therefore, in the present study we use the presence of vertebral deformities as an indicator of decreased strength. Since clinical spine radiographs are not suitable for the determination of texture features, we use hand radiographs for this purpose. In the present study, we investigate whether texture features, measured on hand radiographs, can indicate the strength of vertebral bone, independent of bone mineral density and age.

2. Material and Methods

Material
In the Rotterdam Study[12], a large epidemiological study of the determinants of disease and disability in the elderly, radiographs were taken of, inter alia, the spine and the hands. Using lateral radiographs, vertebral heights were measured and ratios of these heights were calculated according to Melton et al.[13]. If one of the ratios was smaller than a reference value minus 2 SD, a deformity of the vertebra was recorded[14]. We included 118 females with one or more vertebral deformities. The average age was 75.78, ranging from 57 to 93 years. Also, 87 females with no vertebral deformities were included as controls. Of this group the average age was 72.14 years, ranging from 55 to 93 years. The bone mineral density was measured at the femoral neck by dual-energy X-ray absorptiometry.

Texture Features
In each hand-radiograph, a region of interest of approximately 6 mm x 6 mm was selected centrally in the proximal part of the phalanx proximalis of the right index finger. This region was digitized using a CCD-camera; the resulting pixel size was 37.5 μm x 37.5 μm. In this study, we used two texture analysis methods: the Spatial Gray-Level Dependence Method[15] and the Morphological Gradient Method; both methods are described in Chapter 2.

Spatial Gray-Level Dependence Method
In this method, the first step consists of a histogram equalization in which the number of gray-levels is reduced. Thereafter, the co-occurrence features are determined at small scales of 1 to 4 pixels. However, histogram equalization has one drawback: when a background trend is present, this trend is enhanced at the expense of the more subtle gray-level variations at a smaller, spatial, scale. Performing a background correction ensures that, instead of the background trend,

the smaller gray-level variations are quantified. We performed a simple background correction by using a large uniform filter (61x61 pixels).

The histogram of each image was equalized to four gray-levels. Then, for the computation of the intermediate matrix, three distances d were selected: values of 1, 2, and 4 pixels were used. For the angle θ, values of 0° and 90° were taken; the direction 0° was chosen in line with the scan direction of the CCD-camera. As the last step, the five features, as described in Chapter 2, were computed for each combination of distance d and angle θ. An overview of the features used is given in Table 1. These features will be referred to as the co-occurrence features.

Morphological Gradient Method
This method was developed by ourselves, especially for use in radiographs. The features are based on combinations of the basic operations of Mathematical Morphology[16]. Two morphological gradients are used: the G-gradient and the H-gradient. A complete description is given in Chapter 2. The features, which will be referred to as the morphological features, are listed in Table 1. The features HR and GR are sensitive to the roundness of structures and edges, respectively. The features HSS and HSR assess the size of structures and are derived with respectively square and round structuring elements. GSS and GSR quantify the same properties for edges. The orientation of structures and edges present in the image are quantified by HO and GO.

Texture is constituted by the coherence, not the absolute values, of gray-levels. Therefore, texture features should be independent of first order statistics such as the average gray-level and the variance of the gray-levels. An ideal texture analysis method is invariant under strict monotonic gray-level transformations, but except for the Spatial Gray-Level Dependence Method, no texture analysis method meets this requirement. Therefore, this requirement is relaxed to invariance under linear transformations of the gray-levels. The features generated by both methods are independent of first-order gray-level statistics. Whereas the SGLD Method obtains its independence by the histogram equalization step, the features of the MG Method are inherently invariant under linear gray-scale transformations.

143

Table 1. *Texture features generated by the two texture analysis methods.*

Method	Features
SGLDM	$ASM_{d,\theta}$, $CON_{d,\theta}$, $COR_{d,\theta}$, $DMO1_{d,\theta}$, $ENT_{d,\theta}$ for $(d,\theta)=(1,0°),(2,0°),(4,0°),(1,90°),(2,90°),(4,90°)$
MGM	GR_5, GR_7, GR_9, GR_{11}
	$GSS_{7,3}$, $GSS_{9,3}$, $GSS_{11,5}$
	$GSR_{7,3}$, $GSR_{9,3}$, $GSR_{11,5}$
	$GO_{7,1}$, $GO_{9,3}$, $GO_{11,3}$, $GO_{11,5}$, $GO_{13,5}$
	HR_5, HR_7, HR_9, HR_{11}
	$HSS_{7,3}$, $HSS_{9,3}$, $HSS_{11,5}$
	$HSR_{7,3}$, $HSR_{9,3}$, $HSR_{11,5}$
	$HO_{7,1}$, $HO_{9,3}$, $HO_{11,3}$, $HO_{11,5}$, $HO_{13,5}$

Statistical Analysis

The Student's t-test was used to asses whether the texture features differed between the group of females with vertebral deformities and the control group. To test whether the association between a texture feature and the existence of one or more vertebral deformities was related to age or BMD, an age and BMD-adjusted logistic regression analysis was performed, for only those features which showed a significant difference between the groups with and without vertebral deformities ($p<0.05$). The odds ratios are expressed in a standardized form for a 1 SD change in the value of the texture feature.

3. Results

Age was significantly different between the group with vertebral deformities and the control group ($p=0.001$); the difference in the femoral neck density, measured by DEXA, was not significant.

Table 2. *Differences in morphologic texture features between women with vertebral deformities and a control group. The p-values are based on the student's t-test.*

Texture feature	p	Texture feature	p	Texture feature	p
GR_5	<0.01	$GO_{7,1}$	N.S.	$HSS_{9,3}$	<0.05
GR_7	<0.05	$GO_{9,3}$	N.S.	$HSS_{11,5}$	<0.01
GR_9	<0.05	$GO_{11,3}$	N.S.	$HSR_{7,3}$	N.S.
GR_{11}	N.S.	$GO_{11,5}$	N.S.	$HSR_{9,3}$	N.S.
$GSS_{7,3}$	<0.01	$GO_{13,5}$	N.S.	$HSR_{11,5}$	N.S.
$GSS_{9,3}$	<0.01	HR_5	<0.05	$HO_{7,1}$	<0.05
$GSS_{11,5}$	<0.01	HR_7	<0.05	$HO_{9,3}$	N.S.
$GSR_{7,3}$	<0.01	HR_9	<0.05	$HO_{11,3}$	N.S.
$GSR_{9,3}$	<0.01	HR_{11}	<0.05	$HO_{11,5}$	N.S.
$GSR_{11,5}$	<0.05	$HSS_{7,3}$	<0.05	$HO_{13,5}$	N.S.

None of the co-occurrence features showed a significant ($p<0.05$) difference between the two groups. The results for the morphological features are presented in Table 2. The *GR*, *GSS*, *GSR*, *HR*, and *HSS* features are, almost all, significantly different between the two groups, in contrast to the *HSR* features. It can be seen that of the orientation features (*GO* and *HO*) only one is significantly different between the groups with and without vertebral deformities.

In order to investigate whether the differences in texture features are related to age or BMD-differences, a multivariate logistic regression was performed, adjusting for age and BMD. The results are given in Table 3. Some of the odds ratios are, after adjusting for age and BMD, no longer significant. This applies especially to the *HR* and the *HSS* group of features. However, the odds ratios of the *GR*, the *GSS*, and the *GSR* features are still significant after adjusting for age and BMD.

Table 3. *Morphologic gradient features as predictors of vertebral deformities based on logistic regression models.*

Texture feature	Odds ratios (95% CI)	
	Excluding BMD and age	Including BMD and age
GR_5	1.5 (1.1, 2.0)	1.4 (1.1, 2.0)
GR_7	1.4 (1.1, 1.9)	1.4 (1.0, 1.9)
GR_9	1.4 (1.0, 1.9)	1.4 (1.0, 1.9)
GR_{11}	N.S.	N.S.
$GSS_{7,3}$	1.5 (1.1, 2.0)	1.4 (1.0, 1.9)
$GSS_{9,3}$	1.5 (1.1, 2.0)	1.4 (1.1, 2.0)
$GSS_{11,5}$	1.4 (1.1, 1.9)	1.4 (1.0, 1.9)
$GSR_{7,3}$	1.5 (1.1, 2.0)	1.4 (1.0, 1.9)
$GSR_{9,3}$	1.5 (1.1, 2.0)	1.4 (1.0, 1.9)
$GSR_{11,5}$	1.5 (1.1, 2.0)	1.4 (1.0, 1.9)
HR_5	1.3 (1.0, 1.8)	N.S.
HR_7	1.3 (1.0, 1.8)	N.S.
HR_9	1.4 (1.0, 1.9)	1.4 (1.0, 1.9)
HR_{11}	1.3 (1.0, 1.8)	N.S.
$HSS_{7,3}$	1.4 (1.0, 1.8)	N.S
$HSS_{9,3}$	1.4 (1.0, 1.9)	N.S.
$HSS_{11,5}$	1.5 (1.1, 2.0)	1.4 (1.0, 1.9)
$HO_{7,1}$	1.3 (1.0, 1.8)	N.S.

4. Discussion

The strength of bone is not only determined by bone mineral density, but also by the structure of bone. Several authors have shown that texture analysis methods can be used to quantify the structure of trabecular bone in radiographs[1-10]. In this study, we used two texture analysis methods: the Spatial Gray-Level Dependence Method, and the Morphological Gradient Method, the latter method having been especially developed by us for the purpose of quantifying trabecular bone structure in radiographs. We investigated whether texture features generated by these methods, measured on hand radiographs, could indicate the strength of vertebral bone, independent of bone mineral density and age.

Several authors have reported a good discriminative performance of the co-occurrence features on non-radiographic images[15,17,18]. Also when used on mammograms for the detection of breast masses, the co-occurrence features have shown to be useful[19-21]. We found in a previous study, described in Chapter 7, that some co-occurrence features differed significantly ($p<0.05$) between the hand-radiographs of osteoporotic and non osteoporotic females which were matched using the density measured on the same hand-radiograph. In the present study, however, we found that none of the co-occurrence features used was significantly different between the group of females with vertebral deformities and the group without such deformities.

In the study described in Chapter 7, we found also that most of the morphologic gradient features differed significantly between osteoporotic and non-osteoporotic trabecular bone. In the present study, these results are confirmed. However, it should be noted that the majority of the morphologic gradient features which showed a relatively lower significance level in the previous study, were not significantly different in the present study. For example, only one of the orientation features ($HO_{7,1}$) was significantly different between the two groups, whereas in the previous study all orientation features showed a significant difference. A number of factors can be held responsible for the differences in results between these two studies. In the first study, the differences in texture features were studied using a fairly limited number of osteoporotic and non-osteoporotic females who were matched for density as measured in handradiographs. In the present study a much larger population, with and without vertebral deformities, which differed in density was studied. Furthermore, the screen-film combinations used in both studies were not the same.

In order to test whether the differences in texture features were caused by differences in age or BMD, a logistic regression was performed, adjusting for age

and BMD. The odds ratios of the *GR*, *GSS*, and *GSR* features were still significant after adjusting for age and BMD, whereas the odds ratios of most of the *HR* and *HSS* features became non-significant after the adjustment.

In a previous study[10], in which we investigated the correlation between texture features and the strength of vertebrae, we found that in female vertebrae the morphological features and some of the co-occurrence features were correlated with fracture stress. Fracture stress was defined as the ratio of fracture force and the area of the vertebrae. However, when a correction for BMD was performed the co-occurrence features were no longer significantly correlated, in contrast to the morphological features. Taking into account that for the aforementioned study we used magnification radiographs, whereas in the present study a mammographic screen-film system was used, we can state that a selection of morphological gradient features show a relation with vertebral strength, irrespective of film type used.

In this study, we showed that morphological texture features, measured in the proximal part of the proximal phalanx of digitus II, are significantly related, independent of age and BMD, to the presence of one or more vertebral deformities.

Acknowledgements

References

1. Heuck FHW, Bloss WH, Saackel LR, Reinhardt ER. Strukturanalyse des knochens aus Röntgenbildern. Biomedizinische Technik 1980;25:35-42.
2. Lynch JA, Hawkes DJ, Buckland-Wright JC. Analysis of texture in macroradiographs of osteoarthritic knees using the fractal signature. Phys Med Biol 1991;36(6):709-722.
3. Geraets WGM, van der Stelt PF, Netelenbos CJ, Elders PJM. A new method for automatic recognition of the radiographic trabecular pattern. J Bone Miner Res 1990;5(3):227-233.
4. Ruttimann ES, Webber RL, Hazelrig JB. Fractal dimension from radiographs of peridental alveolar bone. Oral Surg Oral Med Oral Path 1992;74:98-110.
5. Caligiuri P, Giger ML, Favus MJ, Jia H, Doi K, Dixon LB. Computerized radiographic analysis of osteoporosis: preliminary evaluation. Radiology 1993;186:471-474.
6. Caligiuri P, Giger ML, Favus MJ. Multifractal radiographic analysis of osteoporosis. Med Phys 1994;21(4):503-508.

7. Benhamou CL, Lespessailles E, Jacquet G, et al. Fractal organization of trabecular bone images on calcaneus radiographs. J Bone Miner Res 1994;9(12):1909-1918.
8. Buckland-Wright JC, Lynch JA, Rymer J, Fogelman I. Fractal signature analysis of macroradiographs measures trabecular organization in lumbar vertebrae of postmenopausal women. Calcif Tissue Int 1994;54:106-112.
9. Caldwell CB, Willett K, Cuncins AV, Hearn TC. Characterization of vertebral strength using digital radiographic analysis of bone structure. Med Phys 1995;22(5):611-615.
10. Veenland JF, Link TM, Konermann W, Meier N, Grashuis JL, Gelsema ES. Unraveling the role of structure and density in determining vertebral bone strength. Calcif Tissue Int 1997;61(6):474-479.
11. Geraets WGM, van der Stelt PF, Elders PJM. The radiographic trabecular bone pattern during menopause. Bone 1993;14:859-864.
12. Hofman A, Grobbee DE, De Jong PTVM, Van den Ouweland FA. Determinants of disease and disability in the elderly: The Rotterdam Elderly Study. Eur J Epidemiol 1991;7:403-422.
13. Melton LJI, Kan SH, Frye MA, Wahner HW, O'Fallon WM, Riggs BL. Epidemiology of vertebral fractures in women. Am J Epidemiol 1989;5:1000-1011.
14. Burger H, Van Daele PLA, Algra D, et al. Vertebral deformities as predictors of non-vertebral fractures. Br Med J 1994;309:991-992.
15. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern 1973;3(6):610-621.
16. Serra J. Image analysis and mathematical morphology. (1st ed.) London: Academic Press, 1982.
17. Weszka JS, Dyer CR, Rosenfeld A. A comparative study of texture measures for terrain classification. IEEE Trans Syst Man Cybern 1976;6:269-285.
18. Ohanian PP, Dubes RC. Performance evaluation for four classes of textural features. Pattern Recognition 1992;25(8):819-833.
19. Chan H-P, Wei D, Helvie MA, et al. Computer-aided classification of mammographic masses and normal tissue: linear discriminant analysis in texture feature space. Phys Med Biol 1995;40:857-876.
20. Petrick N, Chan H-P, Wei D, Sahiner B, Helvie MA, Adler DD. Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification. Med Phys 1996;23(10):1685-1696.
21. Wei D, Chan H-P, Petrick N, et al. False-positive reduction technique for detection of masses ondigital mammograms: global and local multiresolution texture analysis. Med Phys 1997;24(6):903-914.

Chapter 9

General Discussion

Osteoporosis is defined as "a disease characterized by low bone mass and microarchitectural changes of bone tissue, leading to enhanced bone fragility and a consequent increase in fracture risk." (WHO, 1994)[1]. Until recently, the bone mineral density was regarded as the main characteristic of osteoporosis. However, several studies have shown that bone mass is not the only determinant of fracture risk. For example, a large overlap between fracture and non-fracture subjects exists[2-4] and the severity of vertebral fractures correlates poorly with bone mass. Therefore, other factors are thought to play an important role in determining bone strength: one of them is the trabecular bone structure on which we concentrate in this study.

The osteoporotic changes occurring at different skeletal sites have been described using techniques from histomorphometry. In the last decade, less invasive techniques such as CT[5-10] or MRI[11-13] have been used to quantify the trabecular bone structure. In this study we have focussed on quantifying the changes due to osteoporosis in radiographs. Plain radiographs have the advantage that they are easy to obtain, relatively inexpensive and therefore suitable for screening purposes. A drawback of using radiographs for the quantification of structure is that the exact relationship between the three-dimensional bone structure and the two-dimensional projection, appearing in the radiograph as texture, is not known.

Several authors have described the osteoporotic changes in the radiographic trabecular texture[14,15], other authors followed a semiquantitative approach[16-18] by visually classifying the structure in a number of categories. The indices, resulting from such procedures, have a number of drawbacks: the intra- and inter-observer variability tends to be high, only large differences in texture can be detected, and structure and density are combined in one parameter. Since the radiographic density is not only determined by the bone mineral density, but also by the amount of overlying tissue, the exposure conditions, and the development parameters, the categorization may vary with these factors. Computerized texture analysis can overcome most of the drawbacks mentioned: the structure can be evaluated independently of the radiographic density, a more objective and reproducible result can be realized than with human visual inspection, and different aspects of the structure can be evaluated on continuous scales.

In image analysis literature, various texture analysis methods have been described. The majority of these methods is based on an image model, in which certain characteristics of the image texture are condensed. Using such image models, features can be derived which objectively quantify, aspects of, the image texture. Although most of these texture features cannot readily be related to visual

image properties, the features can be used for the purpose of objective characterization of image texture and are therefore useful for discrimination purposes. Since the early seventies, texture analysis methods have been applied to radiographs, for example, to categorize different stages of pulmonary disease[19-22], to detect calcifications and masses in mammograms[23-28] and to quantify bone structure[29-42]. However, most texture analysis methods have not been specifically developed for use on radiographs. Radiographs differ essentially from other imagery, in the sense that they are dose limited: for patient safety, the dose is kept as low as possible. This results in two degrading effects on image quality: noise and image blur, caused by the intensifying screen.

The aim of this thesis was to select or develop texture features which are able to quantify the changes in the radiographic trabecular pattern occurring in osteoporosis. We formulated a number of requirements, listed in Chapter 1, which a texture feature should meet in order to be qualified for the purpose mentioned. These requirements are ordered into three levels. At the first level, requirements are described which are relevant for any texture analysis application. The second level encompasses two requirements which are specifically applicable when using texture analysis in radiographs, whereas the requirements at the third level are solely formulated for the purpose of quantifying the trabecular bone pattern in osteoporosis. In the following, we will discuss, referring to each of the requirements, the outcome of the studies which were partaken to investigate whether the selected texture analysis methods fulfilled these criteria. A summary of the main outcomes is presented in Table 1.

In this study, we focussed on four different texture analysis methods: the Spatial Gray-Level Dependence Method (SGLDM), the Fourier Method (FM), several fractal dimension estimation methods (FD), and the Morphological Gradient Method (MGM). The first two methods, the SGLDM and the FM, have been selected from texture analysis literature, based on their superior discriminative performance on, e.g., aerial photographs. Since the fractal dimension has recently become a popular feature for quantifying texture in radiographs, different fractal dimension estimation methods were combined to form the third method. The fourth method, based on morphological gradients, was developed by us especially for the purpose of quantifying trabecular texture in bone radiographs.

Table 1. *Fulfilment of the requirements listed in Chapter 1 by the different texture analysis methods (n.i. = not investigated).*

			Texture Analysis Methods			
		Requirements	SGLDM	FM	FD	MGM
Texture analysis	1	invariance under monotonic gray-level transformations	+	–/+	–	–
		invariance under linear gray-level transformations	+	+/–	+	+
	2	discrimination	+	+	+	++
Texture analysis	3	independence of noise	+/–	+/–	+/–	+/–
in radio-graphs		independence of MTF	+/–	–	–	+/–
	4	reproducibility	+	–	+	+/–
Texture analysis	5	correlation with bone strength	+	n.i.	n.i.	+
in osteo-porosis		correlation with bone strength independent of BMD	–	n.i.	n.i.	+
	6	discrimination of osteoporotic and non osteoporotic bone	–/+	n.i.	n.i.	+

Texture analysis

Two requirements were formulated which texture features, in general, should fulfil:

Texture features should be invariant under strict monotonic gray-level trans-
formations   *(req. 1)*

Texture features should be able to discriminate between different textures
                                                                        (req. 2)

Invariance under monotonic gray-level transformations
Texture features quantify the spatial pattern laid out by the different gray-levels in the image: texture information is hidden in the coherence of gray-levels, and not in

155

the absolute values of the occurring gray-levels. Therefore, texture features should be able to characterize texture information independent of the absolute gray-level values. Although this is a general requirement, it is especially important when using texture features to quantify texture information present in the radiograph. It is generally acknowledged that diagnosis on radiographs should be invariant under differences in exposure time, processing conditions of radiographic films or differences in scanner settings in digitizing the radiographs. The same applies to texture features. Haralick[43] stated already that texture features should be invariant under monotonic gray-tone transformations. Although this may seem self-evident, several authors have used texture features which are dependent on the first-order statistics of the gray-level distribution, and therefore on the optical density of the radiographs[34]. By performing a background correction, the influence of the average intensity value may be eliminated. However, since the detection of Röntgen quanta is governed by a Poisson process, the variance of the intensity values is still dependent on the average intensity level. Therefore, the results will still be dependent on linear gray-scale transformations. Haralick proposed a histogram equalization step to eliminate the dependence on monotonic gray-scale transformations. Conners[44] tested this procedure in radiographs for normalization of the intensity values for differences in exposure time, development temperature and time, and scanner settings and reported good results.

In this thesis, we used four texture analysis methods: the Spatial Gray-Level Dependence Method (SGLDM), the Fourier Method (FM), Fractal Dimensions (FD), and the Morphology Gradient Method (MGM). We have shown in Chapter 2, that only the features based on the Spatial Gray-Level Dependence Method proposed by Haralick are invariant under monotonic gray-scale transformations, due to the histogram equalization. The features based on the Fractal Method and the Morphology Gradient Method are invariant under linear gray-scale transformations. The features derived from the Fourier Method differ in their dependence on gray-level transformations, varying from invariance under monotonic gray-level transformations, to dependence on the mean and variance of the gray-level distribution.

Therefore, the first step in choosing a texture analysis method for discrimination, categorization or segmentation of different textures, should be a histogram analysis. If only the mean and the variance of the distributions are different, texture analysis methods which are invariant under linear gray-scale transformations can be used. When the gray-level distributions differ in skewness or kurtosis, only texture analysis methods which are invariant under strict monotic

gray-level transformation, such as the Spatial Gray-Level Dependence Method, should be used.

Power of discrimination between different textures
Texture analysis methods have been applied to all kinds of imagery, ranging from aerial and satellite images to medical images, among which radiographs. Most methods have been tested on textures which are fairly distinct from each other, e.g., the Brodatz textures. However, texture analysis methods, when applied to radiographs in order to quantify changes in texture due to a certain disease or condition, should be, in order to be clinically useful, able to detect also gradual changes in texture. The expression of texture changes due to a disease or condition, e.g., osteoporosis, in radiologic images are usually part of a continuum. Whereas human observers, when the texture differences are gradual, can discriminate between a limited number of texture categories only, a computerized texture analysis method should be able to quantify all the different aspects of texture on a continuous scale.

In our study, we used fractal images to test whether the texture analysis methods were able to detect gradual texture differences. We found that all methods used could discriminate fairly well between the different textures, however the Morphological Gradient Method performed superior to the other methods.

Texture analysis in radiographs

In this thesis, we focussed on the application of texture analysis methods to radiographs. Radiographs differ essentially from other imagery in two aspects. In the first place they are dose limited: for patient safety, the exposure dose is kept as low as possible. This has two degrading effects on image quality, namely noise and image blur. The most important source of noise in radiographs, quantum mottle, is the result of statistical fluctuations in the number of x-ray quanta per unit area absorbed by the intensifying screen. The image blur is caused by the intensifying screens which are used in front of the film in order to enhance the light output. The photons, generated in the intensifying screen by the absorption of x-ray quanta, are scattered in the screen before being absorbed by the film. Due to this scattering process, a point in the object is depicted as a blob in the image. This blurring effect results in a reduced spatial resolution, contrast and sharpness. In the spatial frequency domain, this scattering effect is described by the Modulation Transfer Function (MTF). The MTF quantifies the relative transmittance of the spatial

frequencies occurring in the depicted object. Consequently, noise and image blur are both inherently present in radiographs. Second, a clinical radiograph is the result of the projection of three-dimensional anatomical structures. When a three-dimensional structure is non-homogenous, such as trabecular bone, small rotations, e.g., due to repositioning can already result in differences in the structures visible in the radiograph.

Two requirements were formulated which texture features should fulfil in order to be suitable for use in radiographs:

The discriminative power of texture features should be relatively independent of noise and the modulation transfer function   *(req. 3)*

Texture features should be reproducible under repositioning, including small rotations   *(req. 4)*

Dependence on noise and the MTF
The influence of noise and MTF on the discriminative performance of the four texture analysis methods used was evaluated. For that purpose, we simulated the effect of different noise levels and the effect of two different MTF's, corresponding with different screen-film combinations, on fractal images. It appeared that all methods were comparably, negatively, influenced by noise. The effect of the MTF corresponding with the sharp screen-film combination was more detrimental to the Fourier Method and the Fractal Method than to the Spatial Gray-Level Dependence Method and the Morphology Gradient Method. The effect of the MTF corresponding with the unsharp screen-film combination was the least detrimental to the MG Method. It can be concluded from these results that the SGLD Method and the MG Method are more suitable for use in radiographs than the other methods. These results are summarized in Table 1. Since all features are more or less affected by noise and image blur caused by the screen-film combination, it should be stressed that texture features computed from different radiographs can be readily compared only when the noise levels and the screen-film combinations used are comparable. Therefore, the exposure conditions, the object-focus distance, and the screen-film combination should be as similar as possible. This may seem straightforward, however, literature shows that sometimes features computed from radiographs, differing in screen-film combination, are compared[45]. Also the digitizing conditions should be as similar as possible: the digitizer modifies the

image signal with its own MTF. This digitizer-MTF can differ in the two planar dimensions.

Reproducibility

It is well known that differences in repositioning, e.g., small rotations are inevitable in a clinical setting. Consequently, it is important, that texture analysis methods, when used on clinical radiographs, are relatively insensitive to texture variations caused by repositioning or small rotations. We investigated the reproducibility of the features of four texture analysis methods, defining reproducibility as the degree of consistency among repeated measurements performed on the same object under conditions that are as close to identical as possible. For this study femora were used because of the known difficulty in exact repositioning of this bone in vivo, as well as in vitro. When texture features are well reproducible in femora, they can be expected to perform at least equally well in other skeletal sites.

It is common practice to use the coefficient of variation for the determination of the reproducibility, also called precision, of measurements. The coefficient of variation is defined as the ratio of the standard deviation of the different measurements and the average value of all measurements. Thus, the coefficient of variation is by definition dependent on the average value of the measurements. Therefore, when comparing the reproducibility of several variables which differ in average value, the coefficient of variation is not a suitable parameter. Consequently, we based our analysis of reproducibility of the various features only on the observed variances. We distinguished two sources of variance in the measurements: the variance between the subjects and the variance due to repositioning. The variance due to repositioning was expressed as a percentage of the total variance present. We found that most of the features of the SGLD Method and the Fractal Method showed a good to excellent reproducibility, whereas of the Fourier Method more than half of the features displayed a moderate or poor reproducibility. The MG Method performed intermediate: the majority of the features showed an excellent reproducibility, some features were poorly reproducible. These results are summarized in Table 1.

Texture analysis in osteoporosis

Parfitt[46] stated that *"the question is not whether architecture is important to strength, which is self evident and not seriously in dispute, but whether its contribution is more than can be predicted from measurements of bone mass or*

159

density, which is more difficult to establish". In osteoporosis, bone density and structure are two concepts which are physically closely intertwined: when the bone structure deteriorates, the bone mineral density will decrease. On the other hand, when the bone mineral density is low, it can be seen in histomorphological studies that the trabecular structure is impaired. Therefore, there is a strong correlation between bone mineral density and structure: each feature which is able to quantify structure will inherently show a correlation with density. Some authors show explicitly that texture features can discriminate between bone with high and bone with low mineral density[33,38] or correlate significantly with BMD[40]. However, the purpose of the use of texture features or structure features is not to replace bone mineral density measurements; bone mineral density has established itself as an important characteristic of skeletal status. The role of structure features should be an auxiliary one. Structure features should reveal information which, combined with bone mineral density can contribute to a better prediction of fracture risk.

Based on the performances of the four texture analysis methods, using the first four requirements as criteria, two texture analysis methods were selected for further research: the Spatial Gray-Level Dependence Method and the Morphology Gradient Method. Although the fractal dimension estimation methods scored better on reproducibility than the Morphology Gradient Method, we gave priority to the fact that the latter method discriminated better between similar textures and showed less dependence on the MTF. Two additional requirements were formulated which texture features should fulfil in order to be suitable for quantifying the radiographic trabecular bone pattern in osteoporosis:

> *Texture features should correlate with bone strength, independent of bone mineral density* *(req. 5a)*

> *Texture features should be able to discriminate between osteoporotic and non-osteoporotic trabecular bone, independent of bone mineral density*
> *(req. 6a)*

Correlation with bone strength
We investigated the correlation of texture features with bone strength using cadaver vertebrae. Vertebrae were chosen for two reasons. In the first place, vertebrae consist mainly of trabecular bone. Secondly, the simulation of the biomechanical fracture conditions is not as complicated as for other osteoporotic fracture prone skeletal sites. The strength of vertebrae is determined by geometry, bone mineral

density, and structure. The contribution of these three factors to strength should be evaluated separately. This may, again, seem straightforward, however, some authors correlate bone mineral density[47] or texture features[39,48] with compressive strength, not taking into account the contribution of the size of the vertebrae to the compressive strength. Not the compressive load but the compressive stress, which is defined as load divided by the cross-sectional area, should be used when investigating the correlation of strength of vertebrae with density or texture features[49]. Furthermore, texture features and bone mineral density are inherently correlated. Therefore, to investigate the additional contribution of texture features to the prediction of strength apart from density, the correlation with fracture stress should be corrected for bone mineral density.

In our study, we used partial correlation techniques to evaluate the separate contributions of bone mineral density and texture features to strength. We found that, although the Spatial Gray-Level Dependence features correlated significantly with fracture stress, a correction for BMD rendered this correlation non-significant. On the other hand, the morphology gradient features showed, after correction for BMD, still a significant correlation with fracture stress. Multiple regression, including BMD and one morphology gradient feature at a time, resulted in a better prediction of fracture stress than BMD alone. These results, obtained using female vertebrae, are summarized in Table 1. For male vertebrae, the correlations of texture features with fracture stress were slightly lower, and after correction for BMD no longer significant. These findings may be explained by the fact that female vertebrae contain more trabecular bone than male vertebrae[50].

Discrimination between osteoporotic and non-osteoporotic trabecular bone
Since osteoporosis is a systemic condition, we can assume that the structural status of trabecular bone at different skeletal sites will be correlated. Therefore, different skeletal sites can be used to evaluate the structural status of the trabecular bone. However, when using texture analysis methods on radiographs for the quantification of trabecular bone, the number of suitable skeletal sites is limited. Clinical radiographs of the spine and femur, although they have been used[34,36], are not very suitable for texture analysis due to low spatial resolution, the (differences in) overprojections, differences in object-focus distance, and the low repositioning reproducibility. It has been shown that radiographs of wrists and hands are suitable for texture analysis[51]. There are no overprojections, the object-focus distance is fairly constant, repositioning reproducibility is high, and, furthermore, there is the possibility to use a slightly higher dose and a sharper screen-film system for a

161

higher spatial resolution. Therefore, we have used in our studies hand radiographs to quantify the radiographic trabecular pattern.

We performed two studies to investigate the suitability of texture features, determined in hand radiographs, to indicate the osteoporotic status of the spine. In one study, the osteoporotic status was evaluated by the occurrence of vertebral fractures, in the other, vertebral deformities were used for this purpose. In the first study, we also matched the females for density as measured in the same hand radiograph using a method based on radiogrammetry. We found that, although some of the SGLDM features were significantly different between osteoporotic and non-osteoporotic bone in the density-match study, these results were not confirmed in the second study, in which we used a larger, not density-matched, population. On the other hand, the majority of the MG features showed a significant difference in both studies. After a correction for BMD, performed in the second study, a considerable number of MGM features still showed a significant difference between osteoporotic and non-osteoporotic trabecular bone. These results indicate that morphologic gradient features, determined in hand radiographs, can indicate the osteoporotic status of the spine, independent of BMD.

It should be mentioned that when a feature shows a statistically significant difference between two groups, this does not directly mean that this feature can be used for discrimination purposes. To be suitable for discrimination, a feature should show relatively large differences between the groups to be separated, and not too much overlap between the groups. We have shown in the described studies, that texture features, especially the morphologic gradient features, determined in hand radiographs are significantly different between osteoporotic and non-osteoporotic females, independent of BMD. However, this is only the first step. We are still a long way from using texture features, in addition to BMD, in discriminating between osteoporotic and non-osteoporotic females, let alone in predicting fracture risk. Whereas texture features can be used to indicate the structural status of the trabecular bone of the skeleton, the actual occurrence of fractures is not only dependent on BMD and structure, but also on various extra-skeletal factors, such as propensity to fall, impaired postural reflexes during falling, insufficient soft tissue over the point of impact to distribute the force of the fall, and environmental hazards[52].

Conclusions

The main conclusion which can be drawn from the studies described in this thesis, is that the Morphology Gradient Method generates features which seem the most suitable for the quantification of the radiographic trabecular pattern in osteoporosis. Features generated by this method are, independent of bone mineral density, significantly related to the strength of bone and to the occurrence of deformities and fractures in the spine.

Future research should be performed at a fundamental level, e.g., relating the three-dimensional structure of trabecular bone to texture as visible in radiographs, and to texture features, and at a clinical level, e.g., investigating which skeletal sites are the most suitable for structure analysis and whether the texture features determined at that site show predictive value for fractures occurring at other skeletal sites.

References

1. WHO studygroup. Assessment of fracture risk and it's application to screening for postmenopausal osteoporosis. WHO 1994.
2. Kimmel DB, Recker RR, Gallagher JC, Vaswani AS, Aloia JF. A comparison of iliac bone histomorphometric data in post-menopausal osteoporotic and normal subjects. Bone and Mineral 1990;11:217-235.
3. Ross PD, Davis JW, Wasnich RD. Bone mass and beyond; risk factors for fractures. Calcif Tissue Int 1993;53:S134-138.
4. Ott SM. When bone mass fails to predict bone failure. Calcif Tissue Int 1993;53(Suppl. 1):S7-13.
5. Feldkamp LA, Goldstein SA, Parfitt AM, Jesion G, Kleereloper M. The direct examination of three-dimensional bone architecture in vitro by computed tomography. J Bone Miner Res 1989;4(1):3-11.
6. Kuhn JL, Goldstein SA, Feldkamp LA, Goulet RW, Jesion G. Evaluation of a microcomputed tomography system to study trabecular bone structure. J Orthop Res 1990;8:833-842.
7. Chevalier F, Laval-Jeantet AM, Lavel-Jeantet M, Bergot C. CT image analysis of the vertebral trabecular network in vivo. Calcif Tissue Int 1992;51:8-13.
8. Majumdar S, Weinstein RS, Prasad RR. Application of fractal geometry techniques to the study of trabecular bone. Med Phys 1993;20(6):1611-1619.
9. Mundiger A, Wiesmeier B, Dinkel E, Helwig A, Beck A, Moenting S. Quantitative image analysis of vertebral body architecture-improved diagnosis in osteoporosis based on high-resolution computed tomography. Br J Radiol 1993;66:209-213.

10. Ito M, Ohki M, Hayashi K, Yamada M, Uetani M, Nakamura T. Trabecular texture analysis of ct images in the relationship with spinal fracture. Radiology 1995;194:55-59.

11. Wehrli FW, Ford JC, Chung H-W, et al. Potential role of nuclear magnetic resonance for the evaluation of trabecular bone quality. Calcif Tissue Int 1993;53:S162-169.

12. Hipp JA, Janusjwicz A, Simmons CA, Snyder BD. Trabecular bone morphology from micro-magnetic resonance imaging. J Bone Miner Res 1996;11(2):286-292.

13. Majumdar S, Genant HK, Grampp S, et al. Correlation of trabecular bone structure with age, bone mineral density, and osteoporotic status: in vivo studies in the distal radius using high resolution magnetic resonance imaging. J Bone Miner Res 1997;12(1):111-118.

14. Resnick D, Niwayama G. Diagnosis of bone and joint disorders. Saunders Company, 1988.

15. Greenfield GB. Radiology of bone diseases. Philadelphia: J.B. Lippincot Company, 1986.

16. Singh M, Riggs BL, Beabout JW, Jowsey J. Femoral trabecular pattern index for evaluation of spinal osteoporosis. Mayo Clin Proc 1973;48:184-189.

17. Cockshott WP, Occleshaw FRCP, Webber C, Walter SD, O'Brien K. Can a calcaneal morphologic index determine the degree of osteoporosis. Skeletal Radiology 1984;12:119-122.

18. Schnitzler CM, Pitchford DGK, Willis EM, Gear KA. Comparison of the radiographic vertebral trabecular pattern with the vertebral fracture prevalence and spinal bone density. Osteoporosis Int 1993;3:293-299.

19. Sutton RN, Hall EL. Texture measures for automatic classification of pulmonary disease. IEEE Trans Comp 1972;C-21(7):667-676.

20. Kruger RP, Thompson WB, Turner AF. Computer diagnosis of pneumoconiosis. IEEE Trans Syst Man Cybern 1974;4:40-49.

21. Tully RJ, Conners RW, Harlow CA, Lodwick GS. Towards computer analysis of pulmonary infiltration. Invest Radiol 1978;13:298-305.

22. Katsuragawa S, Doi K, MacMahon H. Image feature analysis and computer-aided diagnosis in digital radiography: classification of normal and abnormal lungs with interstitial disease in chest images. Med Phys 1989;16(1):38-44.

23. Caldwell CB, Stapleton SJ, Holdsworth DW, et al. Characterisation of mammographic parenchymal pattern by fractal dimension. Phys Med Biol 1990;35:235-247.

24. Dhawan A, Chitre Y, Kaiser-Bonasso C, Moskowitz M. Analysis of mammographic microcalcifications using gray-level image structure features. IEEE Trans Med Imaging 1996;15(3):246-259.

25. Petrick N, Chan H-P, Wei D, Sahiner B, Helvie MA, Adler DD. Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification. Med Phys 1996;23(10):1685-1696.

26. Thiele DL, Kimme-Smith C, Johnson TD, McCombs M, Bassett LW. Using tissue texture surrounding calcification clusters to predict benign vs malignant outcomes. Med Phys 1996;23(4):549-555.

27. Strickland RN. Wavelet transofrms for detecting microcalcifications in mammograms. IEEE Trans Med Imaging 1996;15(2):218-229.

28. Wei D, Chan H-P, Petrick N, et al. False-positive reduction technique for detection of masses ondigital mammograms: global and local multiresolution texture analysis. Med Phys 1997;24(6):903-914.

29. Heuck FHW, Bloss WH, Saackel LR, Reinhardt ER. Strukturanalyse des knochens aus Rontgenbildern. Biomedizinische Technik 1980;25:35-42.

30. Lynch JA, Hawkes DJ, Buckland-Wright JC. Analysis of texture in macroradiographs of osteoarthritic knees using the fractal signature. Phys Med Biol 1991;36(6):709-722.

31. Geraets WGM, van der Stelt PF, Netelenbos CJ, Elders PJM. A new method for automatic recognition of the radiographic trabecular pattern. J Bone Miner Res 1990;5(3):227-233.

32. Ruttimann ES, Webber RL, Hazelrig JB. Fractal dimension from radiographs of peridental alveolar bone. Oral Surg Oral Med Oral Path 1992;74:98-110.

33. Southard KA, Southard TE. Comparison of digitized radiographic alveolar features between 20- and 70-year-old women. Oral Surg Oral Med Oral Path 1992;74:111-117.

34. Caligiuri P, Giger ML, Favus MJ, Jia H, Doi K, Dixon LB. Computerized radiographic analysis of osteoporosis: preliminary evaluation. Radiology 1993;186:471-474.

35. Chen J, Zheng B, Chang Y-H, Shaw CC, Towers JD, Gur D. Fractal analysis of trabecular patterns in projection radiographs. Invest Radiol 1994;29(6):624-629.

36. Caligiuri P, Giger ML, Favus MJ. Multifractal radiographic analysis of osteoporosis. Med Phys 1994;21(4):503-508.

37. Benhamou CL, Lespessailles E, Jacquet G, et al. Fractal organization of trabecular bone images on calcaneus radiographs. J Bone Miner Res 1994;9(12):1909-1918.

38. Buckland-Wright JC, Lynch JA, Rymer J, Fogelman I. Fractal signature analysis of macroradiographs measures trabecular organization in lumbar vertebrae of postmenopausal women. Calcif Tissue Int 1994;54:106-112.

39. Caldwell CB, Willett K, Cuncins AV, Hearn TC. Characterization of vertebral strength using digital radiographic analysis of bone structure. Med Phys 1995;22(5):611-615.

40. Lee RL, Dacre JE, James MF. Image processing assessment of femoral osteopenia. J of Digit Imag 1997;10(3 S 1):218-221.

41. Smyth PP, Adams JE, Whitehouse RW, Taylor CJ. Application of computer texture analysis to the singh index. Br J Radiol 1997;70:242-247.
42. Veenland JF, Link TM, Konermann W, Meier N, Grashuis JL, Gelsema ES. Unraveling the role of structure and density in determining vertebral bone strength. Calcif Tissue Int 1997;61(6):474-479.
43. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern 1973;3(6):610-621.
44. Conners RW, Harlow C. Equal probability quantizing and texture analysis of radiographic images. Computer Graphics and Image Processing 1978;8:447-463.
45. Webber RL, Underhill TE, Horton RA, Dixon RL, Pope TL. Predicting osseous changes in ankle fractures. IEEE Eng Med Biol 1993;12:103-110.
46. Parfitt AM. Implications of architecture for the pathogenesis and prevention of vertebral fracture. Bone 1992;13:S41-47.
47. McBroom RJ, Hayes WC, Edwards WT, Goldberg RP, White III AA. Prediction of vertebral body compressive fracture using quantitative computed tomography. J Bone Joint Surg Am 1985;67(8):1206-1214.
48. Meier N, Brinkhaus HA, Peters PE, Brinkmann P. Digital Bone structure analysis. In: Boehme JM, Rowberg AH, Wolfman NT, ed. Symposium for Computer Assisted Radiology. Winston-Salem: Symposia Foundation, 1994: 670-675.
49. Mosekilde L, Bentzen SM, Ørtoft G, Jørgensen J. The predictive value of quantitative computed tomography for vertebral body compressive strength and ash density. Bone 1989;10:465-470.
50. Nottestad SY, Baumel JJ, Kimmel DB, Recker RR, Heaney RP. The proportion of trabecular bone in human vertebrae. J Bone Miner Res 1987;2(3):221-229.
51. Geraets WGM, van der Stelt PF, Elders PJM. The radiographic trabecular bone pattern during menopause. Bone 1993;14:859-864.
52. Heaney RP. Is there a role for bone quality in fragility fractures? Calcif Tissue Int 1993;53:S3-6.

Chapter 10

Summary

Osteoporosis is defined as "a disease characterized by low bone mass and microarchitectural changes of bone tissue, leading to enhanced bone fragility and a consequent increase in fracture risk." (WHO, 1994). Until recently, bone mineral density (BMD) was regarded as the main determinant of osteoporosis. However, osteoporosis is not only characterized by a decrease in density, but also by changes in the architecture of bone, e.g., the trabecular bone.

The osteoporotic changes occurring in different bones can be described using techniques from histomorphometry, but a less invasive procedure is offered by radiography. Using X-rays, the 3-dimensional structure of the bone is projected in the radiograph, where it appears as a 2-dimensional texture. Several authors have described or tried to categorize visually the osteoporotic changes which appear in the radiographic trabecular texture.

However, texture is an image characteristic which is difficult to grasp. Intuitively people can discriminate between textures if the differences are large enough, but when the differences become smaller, the human observer is not very suitable for an accurate discrimination. In consequence, visual grading of the trabecular patterns has shown to be not very reliable: the intra- and inter observer variabilities tend to be high. Furthermore, since only rather large differences can be detected, the categorization is usually rather coarse. Another drawback of these visually oriented methods is that in the indices used, radiographic density and structure are combined: the contribution of structure and density cannot be appraised separately. Computerized texture analysis methods are able, in contrast to humans, to assess structures independent of the density present in the image. Therefore, using texture analysis methods it is possible to evaluate the structure independently of the radiographic density and to obtain an objective and reproducible result on a continuous scale.

In image analysis literature, various texture analysis methods have been described. The majority of these methods is based on an image model, in which certain characteristics of the image texture are condensed. Using such image models, features can be derived which quantify, aspects of, the image texture objectively. Although most of these texture features cannot readily be related to visual image clues, the features can be used for the purpose of objective characterization of image texture and are therefore useful for discrimination and segmentation of textures in all kinds of imagery.

Since the early seventies, texture analysis methods have also been applied to medical images, e.g., radiographs, ultrasound scans, CT-scans, and MRI images. In radiographs, texture analysis methods have been used, for example, to categorize different stages of pulmonary disease, to detect calcifications and masses in

mammograms and to quantify bone structure. However, most texture analysis methods have not been specifically developed for use on radiographs. Radiographs differ essentially from other imagery, in the sense that they are dose limited: for patient safety, the X-ray dose is kept as low as possible. Intensifying screens are used to enhance the contrast in the radiograph. Two degrading effects on image quality are the result: noise, due to the low dose, and image blur, which is caused by the intensifying screen.

The goal of the research described in this thesis was twofold: in the first place to assess the suitability of different texture analysis methods for use in radiographs, secondly to select or develop texture features which are able to quantify the changes in the radiographic trabecular pattern occurring in osteoporosis.

In Chapter 1 we formulated a number of requirements, which a texture feature should meet in order to be qualified for the purpose mentioned. In Chapter 2 to 8 the studies partaken to investigate whether the selected texture analysis methods fulfil the criteria, are described.

In this thesis, we focussed on four different texture analysis methods: the Spatial Gray-Level Dependence Method, the Fourier Method, the Fractal Method, which consists of several fractal dimension estimation methods, and the Morphological Gradient Method. The first two methods, the Spatial Gray-Level Dependence Method and the Fourier Method, have been selected from texture analysis literature, based on their superior discriminative performance on, e.g., aerial photographs. Since the fractal dimension has recently become a popular feature for quantifying texture in radiographs, different fractal dimension estimation methods were combined to form the third method. We developed the fourth method, based on morphological gradients, especially for the purpose of quantifying trabecular texture in bone radiographs.

These methods and the corresponding texture features are described in detail in Chapter 2. Also in Chapter 2, a theoretical analysis of the different texture analysis methods is performed to determine the dependence of the different texture analysis methods and the corresponding texture features on the optical density in the radiograph.

In Chapter 3 we investigated the influence of noise and image blur on the discriminative power of the four above mentioned texture analysis methods. We used two types of images, Brodatz and fractal, each consisting of five textures.

Whereas the Brodatz images are very dissimilar, the differences between the fractal images are more gradual. We assume that the behaviour of the different texture analysis methods on the fractal images is more or less representative for the performance on radiologic textures.

First, a baseline classification experiment was performed to asses the discriminative power of the four methods and their individual features on the two sets of images. Subsequently, different noise levels, and the effect of different screen-film systems, causing the image blur, were simulated on these images, after which the discriminative performance of the four texture methods and the constituting features was again evaluated.

On the fractal images, the performance of the Morphological Gradient Method was superior to the other methods, on the Brodatz images, the Morphological Gradient Method performed also better than the other methods, but the difference with the other methods was smaller than for the fractal images. The influence of noise is, as expected, dependent on the image type used. The discrimination of more gradually different images, such as the fractal images, is already lowered for relatively low noise levels. In contrast, when the images are more different, only high noise levels decrease the discriminative performance. Of the four texture methods, the performance of the Morphological Gradient Method is the least hampered by image blur.

From the results presented in Chapter 3 it can be concluded that in circumstances which mimic the conditions prevailing in radiographs, the Morphological Gradient Method is superior to the other methods investigated.

The effect of noise and image blur on fractal dimension estimation methods was studied in more detail in Chapter 4. Since images with a known fractal dimension can be generated, these images can be used as a gold standard for the fractal dimension estimation methods. Apart from comparing the outcomes of the different methods, these images were used as a basis for the simulation of the effects of noise and image blur.

Two types of noise were simulated, quantum mottle and film-grain noise. The effect of image blur was simulated using the Modulation Transfer Functions of different screen-film combinations. In this way, we could quantify the influence of noise and image blur directly on the fractal dimensions estimated by several methods.

The results presented in Chapter 4 showed that, in general, the effect of noise is that the fractal dimensions are overestimated, especially for the lower dimensions, whereas blurring results in an underestimation of the dimensions. The magnitude of

these effects depends on the degree of noisiness and the sharpness of the screen-film system used. Therefore, it can be concluded that dimensions estimated from different radiographs can only be reliably compared when the same screen-film type is used under similar exposure conditions.

The reproducibility of texture features is an often neglected, but nevertheless important issue. Since the 3-dimensional structure of trabecular bone is non-homogenous, small rotations, for example due to repositioning, will result in, primarily small, differences in texture in the radiograph. For a texture analysis method to be clinically useful, it should be robust for small texture variations caused by repositioning which inevitably occur in a clinical setting.

In Chapter 5, we investigated the reproducibility of the features of four texture analysis methods applied to the trabecular pattern of cadaver femora. Reproducibility is here defined as the degree of consistency among repeated measurements taken of a given object under conditions that are as close to identical as experimentally possible. For this study femora were used because of the known difficulty in exact repositioning of this bone in vivo, as well as in vitro. When texture features are well reproducible in femora, they can be expected to perform at least equally well in other skeletal sites.

In general, we found that the reproducibility is dependent on the location of the region of interest. More specifically, the results indicated that most of the features of the Spatial Gray-Level Dependence Method and most of the fractal features are well reproducible. For the Fourier features and the morphology gradient features it was shown that a subset of features displayed a high reproducibility.

The correlation of texture features with bone strength was studied in Chapter 6, using cadaver vertebrae. In this in vitro study we used vertebrae for two reasons. In the first place, vertebrae consist mainly of trabecular bone. Secondly, the simulation of the biomechanical fracture conditions is not as complicated as for other osteoporotic fracture prone skeletal sites.

Of each vertebra, a Direct Magnification Radiograph was taken and QCT was performed to determine the BMD. Then the fracture force was measured under conditions simulating the in vivo situation. Since the strength of vertebrae is determined not only by the BMD and the structure of the bone, but also by the size of the vertebrae, the fracture load should be divided by the cross-sectional area of the vertebra. The resulting parameter is called fracture stress.

Texture features based on two methods, the Spatial Gray-Level Dependence Method and the Morphological Gradient Method, were determined in a region of

172

interest located centrally in each magnification radiograph. A number of texture features of both methods showed a significant correlation with fracture stress. Performing a correction for BMD, the contribution of structure, independent of BMD, to strength could be evaluated. Whereas, after this correction, in female vertebrae none of the SGLDM-features maintained a significant correlation with fracture stress, of the Morphologic Gradient Method, several features still correlated significantly with fracture stress.

These results indicate that morphologic gradient features, determined in magnification radiographs, can contribute, independently of BMD, to the prediction of strength of female vertebrae.

In Chapter 7 and Chapter 8, we investigated the suitability of texture features determined in hand radiographs to indicate the osteoporotic status of the spine. In Chapter 7, the osteoporotic status was evaluated by the occurrence of vertebral fractures, in Chapter 8, vertebral deformities were used for this purpose. In both studies hand radiographs were used.

Radiographs of wrists and hands are very suitable for texture analysis: there are no overprojections, the object-focus distance is fairly constant, repositioning reproducibility is high, and furthermore there is the possibility to use a slightly higher dose and a sharper screen-film system for a higher spatial resolution. A region of interest was selected in the proximal part of the phalanx proximalis of the right index finger. The trabecular structure in this region appears visually as homogenous.

In Chapter 7 we determined, on the same radiographs and in the same finger, the bone mineral density using a method based on radiogrammetry. In order to investigate texture differences between osteoporotic and non osteoporotic subjects independent of BMD, a matching of females with comparable BMD was performed. The results, as presented in Chapter 7, showed that some of the features based on the Spatial Gray-Level Dependence Method, and most of the features based on the Morphological Gradient Method were significantly different for the matched pairs.

In Chapter 8, where a larger population was used, the SGLDM-features were no longer significantly different between osteoporotic and non osteoporotic females. However, several MGM-features were significantly different between females with and without vertebral deformities.

These results indicate that morphologic gradient features, determined in hand radiographs, can indicate the osteoporotic status of the spine.

Chapter 9 contains a discussion of the topics investigated in the foregoing chapters. The main conclusion is that of the four texture analysis methods selected, the Morphology Gradient Method seems the most suitable for texture analysis of the radiographic trabecular pattern in osteoporosis.

Chapter 11

Samenvatting

Osteoporose wordt gedefinieerd als "een ziekte gekarakteriseerd door een lage botmassa en micorarchitecturele veranderingen van het botweefsel, wat leidt tot een verhoogde broosheid van het bot en als gevolg daarvan tot een toegenomen fractuurrisico." (WHO,1994). In het algemeen wordt de botdichtheid beschouwd als de belangrijkste determinant van osteoporose. Echter, osteoporose wordt niet enkel gekarakteriseerd door een afname in botdichtheid, maar ook door veranderingen in de architectuur van het bot, bijvoorbeeld van het trabeculaire bot.

De osteoporotische veranderingen in de verschillende bottypes kunnen worden beschreven door gebruik te maken van technieken ontleend aan de histomorphometrie. Een minder invasieve benadering is mogelijk met röntgenologische technieken. Met behulp van röntgenstralen wordt de driedimensionale trabeculaire structuur van het bot afgebeeld als een tweedimensionale textuur. Een aantal auteurs hebben de veranderingen die zichtbaar zijn in deze trabeculaire textuur beschreven of geprobeerd visueel te classificeren.

Een probleem hierbij is dat textuur een moeilijk te categoriseren beeldeigenschap is. Intuïtief kunnen mensen onderscheid maken tussen texturen zolang de verschillen maar groot genoeg zijn. Zodra de verschillen kleiner worden blijkt een menselijke waarnemer niet erg geschikt voor het maken van een correct onderscheid. Het visueel indelen in klasses van verschillende trabeculaire patronen blijkt als gevolg hiervan niet erg betrouwbaar: de inter- en intrawaarnemer variabiliteiten blijken vaak hoog te zijn. Bovendien is de indeling in categorieën, als gevolg van het feit dat alleen grotere textuurverschillen kunnen worden waargenomen, vaak vrij grof. Een ander nadeel van deze visueel georiënteerde methoden is dat de optische densiteit, dit is de zwarting in de röntgenfoto, en de structuur gecombineerd worden in één index: het is niet mogelijk de structuur en de zwarting afzonderlijk te beoordelen. Gecomputeriseerde textuuranalyse methoden zijn in staat om, in tegenstelling tot de mens, structuur te beoordelen onafhankelijk van de zwarting in het beeld. Door textuuranalyse methoden te gebruiken is het dus mogelijk om structuur, onafhankelijk van de optische densiteit in de röntgenfoto, te waarderen en zo een objectief en reproduceerbaar resultaat op een continue schaal te verkrijgen.

In de beeldanalyse literatuur is een veelheid van textuuranalyse methoden beschreven. De meeste van deze methoden zijn gebaseerd op een model, waarin zekere karakteristieken van textuur zijn samengevat. Door gebruik te maken van deze modellen kunnen textuurkenmerken worden afgeleid die aspecten van de aanwezige beeldtextuur objectief kwantificeren. Hoewel de meeste van deze textuurkenmerken niet rechtstreeks gerelateerd kunnen worden aan visuele

beeldeigenschappen, kunnen deze kenmerken goed gebruikt worden voor een objectieve karakterisering van de beeldtextuur en zijn daardoor geschikt voor het maken van onderscheid tussen en het segmenteren van texturen in allerlei verschillende soorten beelden.

Sinds het begin van de jaren 70 zijn verschillende textuuranalyse methoden ook toegepast op medische beelden, bijvoorbeeld röntgenfoto's, ultrasound scans, CT-scans en MRI beelden. In röntgenfoto's zijn textuuranalyse methoden onder andere toegepast om verschillende stadia van longziektes te categoriseren, calcificaties en massa's in mammografieën te detecteren en botstructuur te kwantificeren. De meeste textuuranalyse methoden zijn echter niet speciaal ontwikkeld voor het toepassen in röntgenfoto's. Röntgenfoto's verschillen essentieel van andere typen beelden in de zin dat ze dosis-gelimiteerd zijn: om stralingshygiënische redenen wordt de röntgendosis zo laag mogelijk gehouden, er wordt een versterkingscherm gebruikt om het contrast te verbeteren. Hierdoor zijn er twee nadelige effecten op de beeldkwaliteit waarneembaar: ruis, als gevolg van de lage dosis, en een beeldversmering, die het gevolg is van het versterkingsscherm.

Het doel van het onderzoek dat beschreven is in dit proefschrift is tweeledig: enerzijds het beoordelen van de geschiktheid voor de toepassing in röntgenfoto's van verschillende textuuranalyse methoden, anderzijds het selecteren c.q. ontwikkelen van textuurkenmerken die in staat zijn om veranderingen, die optreden als gevolg van osteoporose in het trabeculaire patroon in röntgenfoto's, te kwantificeren.

In hoofdstuk 1 hebben we een aantal eisen geformuleerd waaraan een textuurkenmerk zou moeten voldoen om geschikt te zijn voor het bovengenoemde doel. In hoofdstuk 2 tot en met 8 worden de verschillende studies beschreven welke zijn uitgevoerd om te onderzoeken of een aantal geselecteerde textuuranalyse methoden aan deze eisen tegemoet komen.

In dit proefschrift hebben we ons gericht op vier textuuranalyse methoden: een methode die de ruimtelijke relatie van grijswaarden kwantificeert (SGLDM), de fourier methode, de fractale methode die bestaat uit een aantal methoden die de fractale dimensie schatten en de morfologische gradiënt methode. De eerste twee methoden, de SGLDM en de fourier methode zijn geselecteerd op basis van hun superieur onderscheidend vermogen op b.v. luchtfoto's uit de textuuranalyse literatuur. Aangezien er de laatste jaren een toegenomen belangstelling is voor het

gebruik van de fractale dimensie als kenmerk om textuur in röntgenfoto's te kwantificeren, hebben we een aantal verschillende fractale dimensie-schattingsmethoden gecombineerd tot een derde methode. De vierde methode werd, door onszelf, speciaal ontwikkeld voor het kwantificeren van trabeculaire botstructuren in röntgenfoto's. Deze methode is gebaseerd op morphologische gradiënten.

Deze methoden en de bijbehorende textuurkenmerken worden in detail beschreven in hoofdstuk 2. Verder bevat hoofdstuk 2 een theoretische analyse van de afhankelijkheid van de verschillende methoden, en de bijbehorende kenmerken, van de optische densiteit in röntgenfoto's.

In hoofdstuk 3 hebben we de invloed van ruis en beeldversmering op het onderscheidend vermogen van de vier voornoemde textuuranalyse methoden onderzocht. Voor dit onderzoek werden twee beeldtypen gebruikt: Brodatz texturen en fractale beelden. Elk beeldtype werd gerepresenteerd door vijf verschillende texturen. Terwijl de Brodatz texturen onderling sterk verschillen, zijn de verschillen tussen de fractale beelden veel geringer. We veronderstellen dat het gedrag van de verschillende textuuranalyse methoden op de fractale beelden min of meer representatief is voor de prestaties op radiologische texturen.

Eerst hebben we een classificatie-experiment met de twee beeldtypes uitgevoerd om het onderscheidend vermogen van de vier methoden en hun individuele kenmerken te kwantificeren. Vervolgens hebben we verschillende ruisniveaus en het effect van verschillende film-schermcombinaties, die de beeldversmering veroorzaken, gesimuleerd op de verschillende texturen van de twee beeldtypes. Hierna werd opnieuw het onderscheidend vermogen van de vier methoden en hun kenmerken bepaald.

De morphologische gradiënt methode presteerde beter dan de andere methoden. Op de fractale beelden was dit effect het meest uitgesproken. De invloed van ruis was, zoals verwacht, afhankelijk van het gebruikte beeldtype. Het onderscheiden van texturen die onderling weinig verschillen, zoals fractale texturen, wordt al bemoeilijkt bij relatief lage ruisnivo's. Daarentegen kunnen texturen die onderling sterk verschillen ook bij hogere ruisnivo's nog onderscheiden worden. De morphologische gradiënt methode is, vergeleken met de andere methoden, minder gevoelig voor beeldversmering.

Op basis van de resultaten gepresenteerd in hoofdstuk 3, kan worden geconcludeerd dat, onder omstandigheden die de condities nabootsen die gelden in röntgenfoto's, de morphologische gradiënt methode een hoger onderscheidend vermogen heeft dan de andere onderzochte methoden.

Het effect van ruis en beeldversmering op een aantal verschillende methoden die gebruikt worden om de fractale dimensie te schatten hebben we, in detail, bestudeerd in hoofdstuk 4. Hiertoe hebben we een aantal gesynthetiseerde beelden met een bekende fractale dimensie als gouden standaard gebruikt. Naast het vergelijken van de uitkomsten van de verschillende schattingsmethoden, werden deze beelden ook gebruikt als basis voor het simuleren van ruis en beeldversmering.

Twee soorten ruis werden gesimuleerd, quantum en film-korrel ruis. Het effect van de beeldversmering werd gesimuleerd door gebruik te maken van de modulatie-overdrachtfunctie van verschillende film-schermcombinaties. Door deze opzet waren we in staat direct de invloed van ruis en beeldversmering op de fractale dimensies, zoals geschat met behulp van verschillende methoden, te bepalen.

De resultaten, beschreven in hoofdstuk 4, laten zien dat de aanwezigheid van ruis in het algemeen tot gevolg heeft dat de fractale dimensie te hoog geschat wordt. Dit effect is sterker voor de lagere fractale dimensies. Daarentegen resulteert beeldversmering in een te lage schatting van de fractale dimensie. De ernst van deze effecten hangt af van de hoeveelheid ruis en de onscherpte van de gebruikte film-schermcombinatie. Op basis van de resultaten kan geconcludeerd worden dat fractale dimensies geschat uit verschillende röntgenfoto's slechts zinvol vergeleken kunnen worden als de film-schermcombinaties vergelijkbaar zijn en als de foto's onder vergelijkbare condities zijn gemaakt.

De reproduceerbaarheid van textuurkenmerken is een belangrijk maar vaak onderbelicht onderwerp. Als gevolg van de driedimensionale, niet homogene structuur van trabeculair bot, kunnen kleine rotaties, bijvoorbeeld als gevolg van het opnieuw positioneren, al resulteren in, weliswaar kleine, verschillen in textuur in de röntgenfoto. Een textuuranalyse methode moet, wil hij klinisch bruikbaar zijn, bestand zijn tegen deze kleine textuurvariaties.

In hoofdstuk 5 hebben we de reproduceerbaarheid van de kenmerken van de vier gebruikte textuuranalyse methoden onderzocht met behulp van het trabeculaire patroon van vrij-geprepareerde dijbeenderen. Reproduceerbaarheid wordt hier gedefinieerd als de mate van consistentie tussen herhaalde metingen die gedaan worden aan een object onder condities die zo gelijk mogelijk worden gehouden. Voor deze studie hebben we gebruik gemaakt van dijbeenderen aangezien hiervan bekend is dat het exact herpositioneren zowel in vitro als in vivo moeilijk is. Wanneer textuurkenmerken een goede reproduceerbaarheid vertonen in dijbeenderen mag verwacht worden dat deze kenmerken in andere delen van het skelet een gelijke, zo niet betere, reproduceerbaarheid laten zien.

Wij vonden in het algemeen dat de reproduceerbaarheid afhankelijk is van de locatie in het bot van het geselecteerde gebied. Meer specifiek vonden we dat de meeste SGLDM- en fractale kenmerken goed reproduceerbaar waren. Van de fourier en morphologische gradiënt kenmerken vertoonden deelverzamelingen een hoge reproduceerbaarheid.

De correlatie van textuurkenmerken met botsterkte hebben we onderzocht in hoofdstuk 6. Voor deze studie hebben we gebruik gemaakt van vrij-geprepareerde wervels en wel om twee redenen. In de eerste plaats bestaan wervels voor het grootste gedeelte uit trabeculair bot. Ten tweede is het simuleren van biomechanische fractuurcondities voor wervels niet zo gecompliceerd als voor de andere delen van het skelet waar osteoporotische fracturen kunnen optreden.

In deze studie werd van elke wervel een vergrotingsröntgenfoto gemaakt en de botdensiteit werd gemeten met behulp van QCT. Daarna werd de kracht gemeten die nodig was om de wervel te breken onder condities die de in vivo situatie simuleerden. Aangezien de sterkte van de wervel niet alleen bepaald wordt door de densiteit en de structuur van het bot maar ook door de grootte van de wervel, hebben we de "breekkracht" gecorrigeerd voor de oppervlakte van de wervel. De resulterende parameter noemen we de "breekdruk".

De textuurkenmerken gebaseerd op twee methoden, de SGLDM en de morphologische gradiënt methode, hebben we bepaald in een regio centraal in elke wervel op de vergrotingsröntgenfoto. Een aantal textuurkenmerken van beide methoden bleek significant gecorreleerd te zijn met de "breekdruk". Door een correctie uit te voeren voor de botdensiteit, kan de bijdrage van de structuur aan de sterkte, onafhankelijk van de densiteit, geëvalueerd worden. In de wervels afkomstig van vrouwen bleek na deze correctie geen van de SGLDM-kenmerken nog significant gecorreleerd met de "breekdruk". Echter, verscheidene kenmerken van de morphologische gradiënt methode vertoonden, na de densiteitscorrectie, wel een significante correlatie met de "breekdruk".

Deze resultaten tonen aan dat morphologische gradiënt kenmerken, bepaald in vergrotingsröntgenfoto's, een bijdrage, onafhankelijk van de botdensiteit, kunnen leveren aan de voorspelling van de sterkte van wervels.

In hoofdstuk 7 en 8 hebben we onderzocht in hoeverre textuurkenmerken die bepaald kunnen worden in hand-röntgenfoto's geschikt zijn als indicatoren voor de osteoporotische toestand van de wervelkolom. Terwijl in hoofdstuk 7 de osteoporotische toestand van de wervelkolom werd geëvalueerd aan de hand van

het aantal wervelfracturen, werd hiervoor in hoofdstuk 8 het aantal wervelvervormingen gehanteerd.

In beide studies hebben we gebruik gemaakt van hand-röntgenfoto's. Hand- en pols-röntgenfoto's zijn zeer geschikt voor textuuranalyse: er zijn geen overprojecties, de object-focus afstand is vrij constant, de reproduceerbaarheid bij het opnieuw positioneren is vrij groot en bovendien kan er een iets hogere dosis gebruikt worden en een scherpere film-schermcombinatie voor een hogere ruimtelijke resolutie. In de rechter wijsvinger, in het proximale deel van het proximale botje, hebben we een regio geselecteerd met een homogene trabeculaire structuur.

In hoofdstuk 7 hebben we met behulp van dezelfde foto en in dezelfde vinger de botdensiteit bepaald gebruik makende van een methode gebaseerd op radiogrammetrie. Om textuurverschillen tussen osteoporotische en niet osteoporotische vrouwen onafhankelijk van de botdensiteit te kunnen onderzoeken zijn de textuurkenmerken van vrouwen uit de twee groepen met overeenkomstige densiteit met elkaar vergeleken. De resultaten, weergegeven in hoofdstuk 7, laten zien dat enkele SGLDM-kenmerken en de meeste morphologische gradiënt kenmerken significant verschillen tussen osteoporotische en niet-osteoporotische vrouwen.

In de studie beschreven in hoofdstuk 8, waarvoor we een grotere populatie hebben gebruikt, verschillen de SGLDM-kenmerken niet significant tussen osteoporotische en niet-osteoporotische vrouwen. Van de kenmerken gebaseerd op de morphologische gradiënt methode zijn echter wel verscheidene kenmerken significant verschillend tussen vrouwen met en zonder wervelvervormingen.

Deze resultaten laten zien dat morphologische gradiënt kenmerken, bepaald in röntgenfoto's van de hand, gebruikt kunnen worden om een indicatie te geven van de osteoporotische toestand van de wervelkolom.

Hoofdstuk 9 bevat de algemene discussie van de onderwerpen die zijn onderzocht in de voorgaande hoofdstukken. De belangrijkste conclusie is dat, van de vier onderzochte textuuranalyse methoden, de morphologische gradiënt methode het meest geschikt lijkt voor textuuranalyse van osteoporotische veranderingen in trabeculair bot in röntgenfoto's.

Dankwoord

Vele mensen hebben op enigerlei wijze bijgedragen aan de totstandkoming van dit proefschrift:

Prof. dr E.S. Gelsema en dr ir J.L. Grashuis, beste Edzard en Jan: vele, soms zeer levendige, discussies, bleken nodig om de nogal weerbarstige materie in een aanvaardbaar begrippenkader te persen. Het resultaat daarvan is neergelegd in dit boekje. Zonder deze kritische inzet van jullie was dit proefschrift niet tot stand gekomen. De open en prettige sfeer waarin we altijd hebben kunnen samenwerken heb ik zeer gewaardeerd. Bovendien heb ik van jullie de ruimte gekregen om mijn eigen weg te zoeken en volgens mijn eigen, steeds uitdijende, planning te werken. Edzard, jij hebt een zeer belangrijke bijdrage geleverd aan de structurering en de begrijpelijkheid van de geschreven artikelen. Ogenschijnlijk onvermoeibaar heb je vele versies gecorrigeerd. Jan, jij was altijd bereid om mee te denken over, en oplossingen te zoeken voor de meer fundamentele problemen die ik in de loop van het onderzoek tegenkwam. Daarnaast kon ik ook altijd met meer praktische zaken, zoals bijvoorbeeld het grafisch presenteren van data, bij jou terecht.

Frits van der Meer: de discussies met jou heb ik altijd als zeer nuttig en aangenaam ervaren. Deze discussies waren onontbeerlijk voor mijn begripsvorming op het terrein van de fysica van het röntgenproces, maar ook het denken in het frequentie-domein heb ik met jouw hulp weten te ontwikkelen. Het samen puzzelen, over bijvoorbeeld de relatie tussen MTF, OTF, PSF en LSF, werkte zeer verhelderend.

Guus Beckers: vooral in het beginstadium van mijn onderzoek heb je een wezenlijke bijdrage geleverd, wat onder andere zijn weerslag heeft gevonden in het artikel over de fractale dimensie schatting. Ook kon ik bij jou aankloppen voor ondersteuning op het gebied van de, voor mij, nog steeds duistere krochten van de unix-machinerie.

I would like to thank Thomas Link, for making available the radiographs, fracture loads and other data of the vertebrae study, and for creating the opportunity for us to make a considerable number of DIMA-radiographs for the reproducibility study.

Rick van Rijn en Wiebeke van Leeuwen wil ik bedanken voor de ondersteuning bij het verzamelen, en ingieten van de femora en het fotograferen hiervan op het DIMA apparaat in Münster.

Andries Zwamborn en Teun Rijsdijk hebben een belangrijke bijdrage geleverd aan de presentatie van mijn onderzoek op verschillende congressen door het leveren van de ondersteuning bij het maken van posters en dia's.

Curriculum Vitae

Jifke Frouwke Veenland was born on May 20, 1965 in Drachten. She attended Gymnasium-β at the Drachtster Lyceum from 1977 till 1983. In 1983 she started her medical studies at the University of Groningen. She obtained her Master of Science degree in 1987, after which she studied Computer Science at the University of Groningen and the University of Leiden. In 1991 she graduated on the project: "Evaluation of Medical Expert Systems". During both studies she was working as a student-assistant teaching in different courses: Medical Physics, Medical Informatics and Statistics, Databases, Programming, Stochastics, and Computer Assisted Medical Decision Making.

In 1991 she started working at the Department of Medical Informatics and the Department of Experimental Radiology at the Erasmus University Rotterdam. During this time, besides working at the studies described in this thesis, she was active in a number of committees: the MAIOR: Medisch AIO-Overleg Rotterdam, the IFAR: Interfacultair AIO-overleg Rotterdam and the LAIOO Landelijke AIO-Overleg.

Since 1997 she is working for KPN Research in Leidschendam at the department Automated Mail Systems, in the field of image analysis, intelligent character recognition, word recognition, and automated form reading.