MARKOV RANDOM FIELDS (MRF)-BASED TEXTURE SEGMENTATION FOR ROAD DETECTION

Ralf Reulke[a], Artur Lippok[b]

[a]Humboldt-University Berlin, Department of Computer Science, Computer Vision
Unter den Linden 6, 10099 Berlin, Germany - reulke@informatik.hu-berlin.de
[b]Hilbertstr. 22a, 12307 Berlin, Germany - lippok@gmx.de

Commission III, WG III/5

KEY WORDS: Mapping, Photogrammetry, Vision, Fusion, Identification, Monitoring, Real-time, Platforms

ABSTRACT:

Traffic observation from airplane platforms using digital cameras is a fairly new application of Video Image Detection Systems (VIDS). These systems are also particularly interesting for observations at major events or in catastrophe situations. Applied systems are small or medium-sized panchromatic cameras. In order to concentrate on road image processing algorithms were applied to distinguish between urban and surrounding vegetation areas. Therefore, texture-based segmentation seems to be an adequate approach, because of the panchromatic images. Attention is focused on MRF, which are examined concerning their use for texture-based extraction of street space in aerial photographs.

KURZFASSUNG:

Der Einsatz von digitalen Kameras für Verkehrsbeobachtungen von Flugzeugplattformen ist eine relativ neue Anwendung von Videodetektionssystemen (VIDS). Diese Systeme sind auch besonders interessant für Überwachungsaufgaben bei Großveranstaltungen oder in Katastrophensituationen. Die angewendeten Systeme sind kleine oder mittelgroße panchromatische Kameras. Bildverarbeitungsalgorithmen zur Unterscheidung von urbanen und umgebenen Vegetationsgebieten wurden angewendet, um die Extraktion von Fahrzeugen lediglich auf den Straßenbereich zu fokussieren. Aufgrund der panchromatischen Bilder erscheint die texturbasierte Segmentierung dafür der adäquate Ansatz. Besonderes Augenmerk gilt den MRF, die hinsichtlich ihrer Nutzung für texturbasierte Extraktion von Straßenraum in Luftbildern untersucht werden.

1. INTRODUCTION

Traffic observation from airplane platforms using digital cameras is a new application of Video Image Detection Systems (VIDS) as already used for regular traffic observations. Monitoring from airplane platforms is also particularly interesting for observations at major events or in catastrophe situations. An overview of current airborne and spaceborne monitoring systems and related activities can be found in (Hinz et al., 2006).

The investigations described here, are based on the ANTAR-system and the evaluation software "Traffic Finder" developed by the German Aerospace Center. The system provides real-time extraction of traffic data. ANTAR includes a digital mega-pixel camera and an inertial measurement unit. The acquired images are transmitted to the ground and analysed by the software to extract cars and their features (Ruhe et al, 2007).
One essential feature of such a system is the immediate knowledge of extracted cars or derived traffic parameters of street sections, which can be annotated in digital street maps (e.g. made by NAVTEQ). To improve the effectiveness of image processing algorithms, irrelevant traffic areas are omitted. In the presented example the observation area is defined with help of digital street maps in combination with a digital elevation model. For that purpose the camera orientation is exactly determined. Unfortunately, the precision of roadway geometry is restricted because of navigation errors, limited map accuracy or map availability. Therefore, the area with occurring

traffic shall be exactly determined using additional image processing algorithms. Utilized camera systems are small or medium-sized panchromatic cameras. Image processing is restricted to texture-based segmentation approaches due to the panchromatic images.. Thus multispectral and colour approaches are not applicable.

Segmentation is a commonly used term for identifying differences between particularly interesting and uninteresting objects as well as distinguishing foreground from background content. In order to identify possibilities for acquisition of scene information by digital images an analysis of the principle features of these images is required. In this regard, textures are the only possibility to derive information from imagery, besides the grey or colour values and structural features. For this reason, texture analysis methods were used since the beginning of digital image processing. There are continuous suggestions for new algorithmic approaches to texture-based scene analyses. Thereby, also Laws-Energy, Fuzzy- and Fourier approaches are analyzed besides first and higher order statistics.

Road detection is a prerequisite for traffic observation from airplane platforms. The paper discusses different approaches to texture-based feature extraction and segmentation. Standard algorithms are examined first. Caused by the complicated observation conditions these methods cannot be used successfully.

Attention is focused on MRF, which are examined concerning their use for texture-based extraction of streets from aerial photographs. Besides the theoretical principles, also the methods and aspects of implementation are discussed such as simulated annealing, parameter analysis according to Maximum Pseudo Likelihood (MPL) and Conditional Least Square (CLS). In addition, different influences on the derivation of texture parameters are researched including the response to noise, scaling as well as to size and configuration of neighbourhood systems. This is referred to familiar textures (Brodatz-textures). In this paper, the application of MRF on aerial images is specified. Evidence is provided regarding the applicability of these approaches to distinguish between urban areas and vegetation.

Afterwards the MRF are successfully applied to real image data. The paper ends with a summary and an outlook.

2. STANDARD APPROACHES

An useful compilation of standard methods can be found in (Tomita & Tsuji, 1990).

Well-known approaches are based on the co-occurrence matrix (see e.g. Haralick, 1973) and derived parameters (e.g. contrast, energy, entropy). Single texture features are unsuitable for texture segmentation, caused by different viewing and illumination conditions as well as shadows, etc. Since single statistical texture features don't permit image segmentation examinations were carried out using couples of texture features. All combinations were examined based on the first order statistics and the statistics were derived from the second order co-occurrence matrix. Only a few feature combinations allow good picture segmentation. Best results could be achieved using the couple of texture features entropy and energy (see figure 1). Segmentation results from single images cannot be generalized. Completely different results are obtained using different images with the same features.
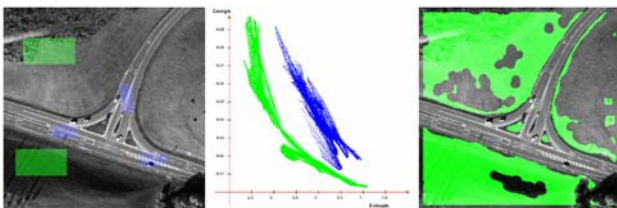

Figure 1. Segmentation results of a single image using entropy and energy. Left: Selection of the training sub-areas vegetation (green) and road (blue). Middle: The entropy over energy diagram. Right: Segmentation result.

Similar results were also achieved by using Laws filter and Fourier approaches with a circle shaped spatial filter. The energies correspond to higher order statistics. Laws has introduced a uniform approach to derive many different texture features from an image. He defines linearly independent base vectors of the size 3, 5 or 7. These base vectors reflect respectively certain qualities such as intensity, edges or waviness. A designated couple is chosen from these base vectors and calculates a square mask using the dyadic product. The image is then convolved by this mask. Figure 2 shows some results of different combinations of couples of Laws energy vectors (Laws, 1980).
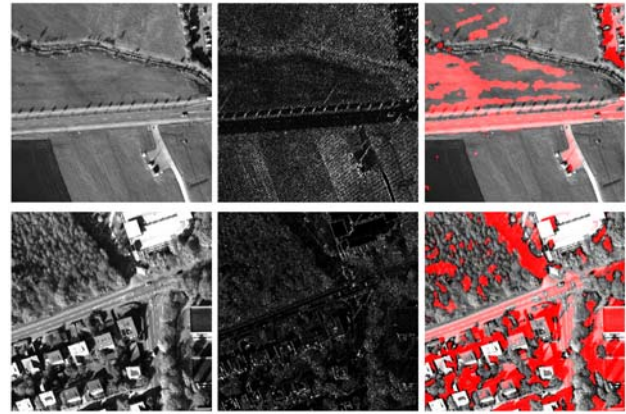

Figure 2: Segmentation results using Laws texture feature. Left: Original image, Middle: texture image, Right: classified image. Top and bottom line show different combinations of couples of Laws energy vectors.

Unfortunately, the known problems appear also at the Laws features. The results vary strongly between the images, shade leads to problems and single objects (such as vehicles) are wrongly classified.

It is to conclude that the classic attempts at texture-based segmentation are not sufficient.

3. RELATED REPRESENTATIONS AND MRF

3.1 MRF-based Segmentation

For segmentation and classification in remote sensing applications, MRF-based methods have been established for several years. The MRF-model offers several advantages for image segmentation and classification. A finite number of parameters characterize spatial interactions of pixels in order to describe an image region.

Texture description and texture generation are essential steps for scene understanding and scene segmentation. Zhao et al. implemented an infrared battlefield simulation to test infrared weapons (Zhao et al., 2007). The MRF-model was used to synthesize various kinds of visible textures. Gaussian-Markov random fields (GMRF) were firstly used to sample the distribution of temperature fields on terrain surfaces.

Images with both fine spectral and spatial sampling (hyperspectral images) can be used to perform a joint texture analysis in both discrete spaces. To achieve this goal a probabilistic vector texture model, using a GMRF was proposed (Rellier et al., 2004). The parameters allow the characterization of different hyperspectral textures. An application of this work is the classification of urban areas.

Descombes et al. focus on GMRF and propose therefore different estimation methods (Descombes et al., 1999). They demonstrated the potential of the estimated parameters to carry out texture differentiation for remote sensing data.

3.2 MRF modelling

The image is called Markovian, if the probability distribution of grey levels for each pixel Z_i merely depends on the grey levels of the neighbouring pixels Z^i. These images can be described by

a Gibbs fields. For this approach the image is characterised by the conditional probability. This probability distribution denotes the probability for accepting a grey level g_i by a pixel Z_i, assuming the neighbouring pixels Z^i have the grey levels g^i (see e.g. Gimel'farb, 1999).

$$P(Z_i = g_i \mid Z^i = g^i) = \frac{1}{S} e^{-H(g_i, g^i)}$$

with   (1)

$$S = \sum_{g=0}^{G} e^{-H(g, g^i)}$$

S is called partition sum, while H stands for the energy function characterizing the Markov random field. In most cases, this energy function is parameterised by a vector $\theta = [b_0, b_{1,\dots}]^T$.

This parameter vector characterizes the texture and can be used for segmentation as a multidimensional texture feature. The used autobinomial model has the following energy function (Schroeder et al., 1998):

$$H_\theta(g) = H_\theta(g_i, g^i) = -\ln \binom{G}{g_i} - g_i \cdot \eta$$

with   (3)

$$\eta = b_0 + \sum_j b_j \frac{g_j + g_j'}{G}$$

G is the maximum grey value. In the first term each grey value g_i of the pixel Z_i is weighted separately. The middle grey levels are stronger weighted than the bright or dark grey levels of a pixel, because of the special structure of the autobinomial coefficients. In the second term the grey values g_i are applied together with the grey values g_i' of their neighbouring pixels.. In doing so different neighbourhood systems (see e.g. Gimel'farb, 1999) can be used (see Figure 3).
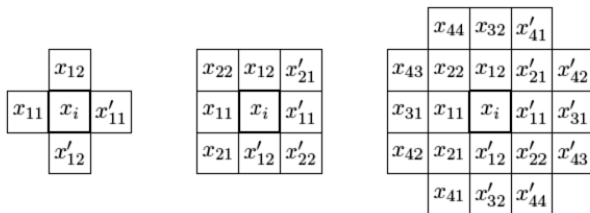

Figure 3. Frequently used neighbouring systems

The parameter vector θ has to be determined in order to segment a given image. For this purpose two different methods were used. The first method is to determine the parameter using the maximum pseudo likelihood (MPL) estimation. Through this approach the overall probability for the image is maximised. Unfortunately, the overall probability for the autobinomial model is unknown. For this reason, Besag's suggestion to approximate the overall probability by the product of all conditional probabilities was used (Besag, 1986):

$$\theta_{MPL} = \arg \max_\theta P_\theta(g)$$

with the approximation   (3)

$$P_\theta(g) \approx \prod_{i \in \Re} P_\theta(g_i \mid g^i)$$

Different local maxima may exist due to the complexity of this probability distribution function. The simulated annealing method was used to avoid local maxima and to find the global maximum.

As second method for parameter determination, the conditional least square (CLS) estimation was used (Schroeder et al.,1998). This approach attempts to minimise the error between the given picture and the expected image. By means of the least square adjustment calculus the parameter vector   can be calculated as follows:

$$\theta_{CLS} = \arg \min_\theta \sum_i (g_i - \tilde{g}_i)^2$$

and   (4)

$$\hat{\theta} = (G^T G)^{-1} G^T d$$

While $\tilde{g}_i$ represent the expected grey values, G represents a matrix containing grey values of neighbouring pixels and d represents a logarithmic vector of these grey values. Both methods achieve the same level of quality, but the calculus of CLS is significantly faster and was therefore used for further calculations.

The investigation of MRFs for texture segmentation was carried out in several steps. Firstly, the validity of the parameter extraction was analysed to examine whether the parameter extraction is reproducible, and to find conditions under which the algorithm is stable. Synthesised MRF images were used in order to compare the extracted with the correct parameters. The images were synthesised using a Markov chain and the Gibbs sampler from a given parameter vector (thus from a given MRF). This iterative procedure is explained in (Paget & Longstaff, 1998). Typical parameter sets were extracted from Brodatz textures (Brodatz, 1966). The effects of noise, image scaling and neighbourhood system size as well as the parameter normalisation were also examined.

In order to achieve a quasi-continuous scaling and to avoid aliasing effects, the bicubic resampling method has been applied. Examinations of the fourier spectrum of the scaled pictures have shown that low-frequency information is preserved by this method.

Having investigated the parameter extraction the segmentation of Brodatz textures has been researched. Using Brodatz textures various influencing factors like the sunset were avoided. The segmentation has been realised using the point-based method of support vector machines. This segmentation system is appropriate, because it facilitates fast segmentation in a multidimensional feature space.

All segmentations were made in the quasi-continuous scaling space for quantifying the influence of scaling under different

conditions. The results were illustrated in a chart with logarithmic scale abscissa.

4. RESULTS AND DISCUSSION

All analyses were performed using different images or image collages. Some examples are shown below (figure 4).
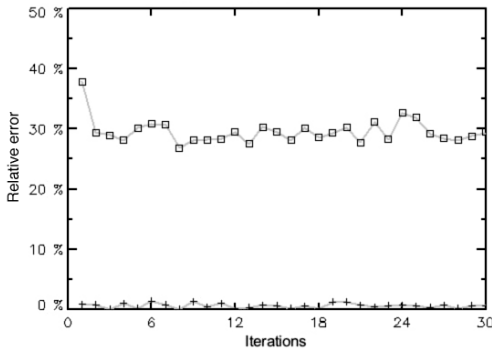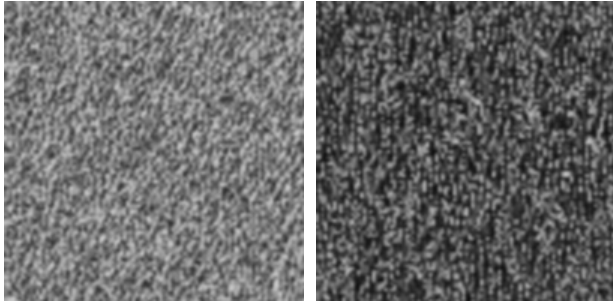

Figure 4. Synthesised textures (top). The upper left picture was used for this calculation; the upper right is the normalized image. Error in parameter extraction (bottom) with (+) and without ( ) parameter normalisation.

4.1 Parameter Validation

Extractions of the parameter vector from synthesised images do not always work perfectly. Although the extracted parameter vector has the right orientation it has the incorrect length, which has to be normalised to unit length to ensure reproducibility of the extraction. This effect can be shown with synthesised textures as seen in Figure 4.

Using the normalised parameters, the extraction can be done with relatively small errors. The parameter extraction is very sensitive to sensor noise. Even if there is already an obvious degradation at a noise level of 5% a normalisation of the parameter vector has a positive effect (see Figure 5).
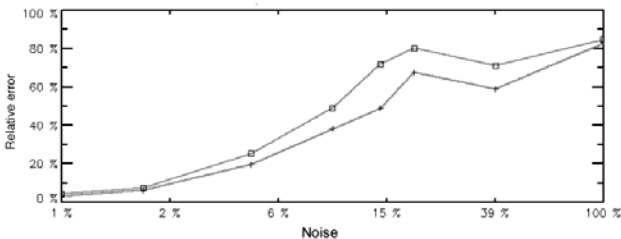

Figure 5. Image segmentation error with different degrees of noise with (+) and without ( ) parameter normalisation.

Neighbourhood system extensions have barely influence on parameter extraction, because with it only new parameters are added. In enlarged neighbourhood systems (NBS) the existing parameters represent the same pixel interaction.

In contrast to that the scaling of the picture has a strong influence on the model parameters.

4.2 Factors that influence segmentation

Radiometric normalisation of the image is the first processing step. Thus, the segmentation is not always improved, but shading effects can be avoided effectively (see Figure 6). The normalisation to a uniform arithmetic mean and deviation leads to better results than a histogram equalization.
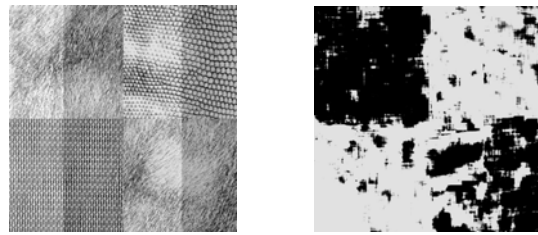

Figure 6a. Segmentation result with shading effects: Original (left), shadow removal without normalisation (error = 16%) (right)
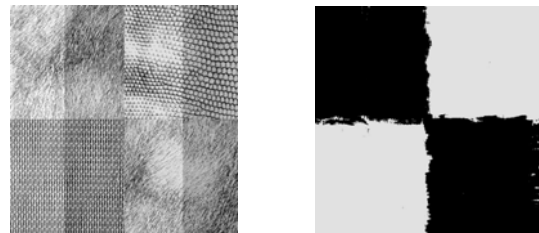

Figure 6b. Segmentation result with shading effects. Original (left), shadow removal with normalisation (error = 2%) (right)

The normalisation of the extracted parameter vector also yields better segmentation results. The example has already indicated that the parameter normalisation is required for valid parameter extraction. Accordingly, this has a positive impact on the segmentation as well.
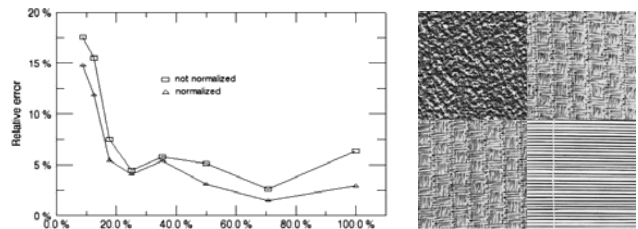

Figure 7. Example of segmentation errors with and without parameter normalisation (left) and the original image (right)

The texture window size and the scaling strongly affect the segmentation (see Figure 7). However, for each picture different values lead to an optimal result. In general the following trend can be derived:

The finer the texture and the higher the scaling the smaller the texture window should be chosen. However, no general rules can be defined from these results. The texture window size and the scaling must be adjusted for every special texture.

The neighbourhood system, by contrast, has only little influence on segmentation results. The reason is that only a few parameters contribute to the indication of texture quality. Additional parameters (a bigger neighbourhood system) do not contain extra information and do not enhance segmentation results. A NBS with 10 neighbours is usually adequate, and therefore was used for further investigations.
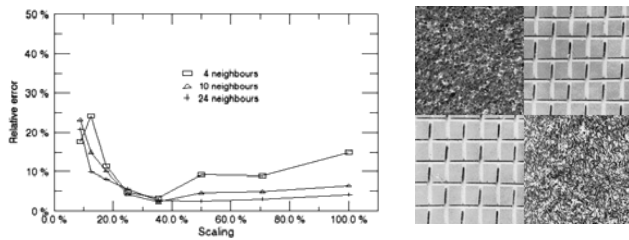

Figure 8: Example of segmentation errors with different neighbourhood systems (left) and the original image (right)

4.3   Results with airborne images

The methods and approaches presented were taken into account for segmentation of airborne images. Images taken by a DMC airborne camera (Intergraph GmbH) were radiometrically normalised, the model parameters were normalised to a uniform length after extraction and a neighbourhood system with 10 neighbours was used. The segmentation was carried out for different texture window sizes and scaling.

Good results could be achieved with the used pictures when the scaling was larger than 50% and the texture windows were small (3-5 pixels). Nevertheless, the segmentation is affected if there is poor image radiometry (see figure 9, middle picture). Moreover, heterogeneous areas, e.g. urban areas with many single trees, make segmentation difficult, though in these cases a crude discrimination is possible.

5.   CONCLUSION AND OUTLOOK

The paper considers road detection in panchromatic images for traffic observation from airplane platforms. Because of the limited image size structure based approaches cannot be applied. Therefore, in this paper texture based algorithms are utilized.

The influence of image quality and pre-processing, image scaling, size of the texture window and size of the neighbourhood system, as well as the influence of parameter normalization where investigated.

In this context it is interesting to realize that the MRF characteristics are independent of illumination of the observed area. According to this it is possible to minimize the influence of cast shadow - a common problem in natural scenes. Altogether, it could be shown that the described method is suitable for the distinction of streets and surrounding areas. For real-time applications the algorithm and the implementation must be optimized.
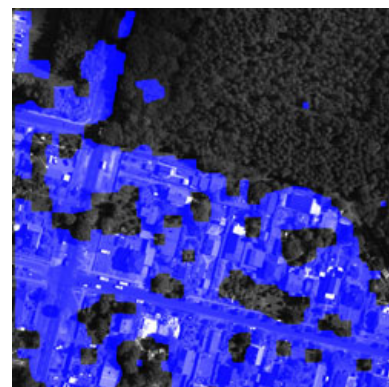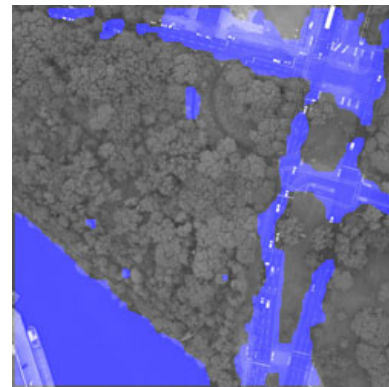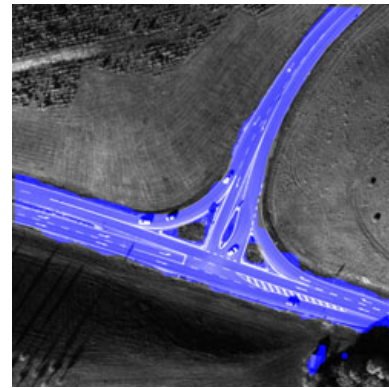

Figure 9. Three examples of segmentation results with aerial photos. Blue is the extracted road area.

REFERENCES

Besag, J., 1986. On the Statistical Analysis of Dirty Pictures, *Journal of the Royal Statistical Society*, B26, pp. 25-36.

Brodatz, P., 1966. A Photographic Album of Artists and Designers, Dover Publications

Descombes, X., Sigelle, M., and Preteux, F.,1999. Estimating Gaussian Markov random field parameters in a nonstationary framework: Application to remote sensing imaging. *IEEE Transactions on Image Processing*, 8(4), pp. 490-503.

Gimel'farb, G., 1999. *Image Textures and Gibbs Random Fields*. Kluwer Academic Publishers, Dordrecht/Boston/ London

Haralick R., Shanmugam K., Dinstein I., 1973. Textural Features for Image Classification. *IEEE SMC* 3, pp. 610-621

Hinz, S., Bamler, R., Stilla, U., 2006. Theme Issue: Airborne and Spaceborne Traffic Monitoring. *ISPRS Journal of Photogrammetry and Remote Sensing*, 61(3-4), pp. 135-280

Laws, I., 1980. Texture Image Segmentation. Ph.D. Thesis, TR. No. 940, Image Processing Institute, University of Southern California

Paget, P., Longstaff, D., 1998. Texture Synthesis via a Noncausal Nonparametric Multiscale Markov Random Field. *IEEE Transactions on Image Processing*, 7(6), pp. 1-9

Rellier, G., Descombes, X., Falzon, F., Zerubia, J., 2004. Texture feature analysis using a Gauss-Markov model hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 42(7), pp. 1543-1551.

Ruhe, M., Kuehne, R., Ernst, I., Zuev, S., Hipp, E., 2007. Air Borne Systems and Datafusion for Traffic Surveillance and Forecast for the Soccer World Cup. TRB 2007 (86. Annual Meeting), Washington, DC (USA), 2007-01

Schroeder, M., Rehrauer, H., Seidel, K., Datcu, M.,1998. Spatial Information Retrieval from Remote-Sensing Images – Part II: Gibbs-Markov Random Fields. *IEEE Transactions on Geoscience and Remote Sensing*, 36(5), pp. 1446–1455.

Tomita, F., Tsuji, S., 1990. *Computer Analysis of Visual Textures*. Kluwer Academic Publishers, Boston/Dordrecht/ London

Zhao, Y. D., Zhang, L. P., Li, P. X., Huang, B., 2007. Classification of high spatial resolution imagery using improved Gaussian Markov random-field-based texture features. *IEEE Transactions on Geosciences and Remote Sensing*, 45(5), pp. 1458-1468.