

## FUSION OF STOCHASTIC PROPERTIES BY FUZZY INTEGRALS AND APPLICATIONS ON DETECTION WITHIN IMAGES

Hartwig HETZHEIM  
Institute of Space Research and Planetary Exploration, DLR, Berlin, Germany  
hartwig.hetzheim@dlr.de

Working Group III/5

**KEY WORDS:** Fusion of properties, image processing, fuzzy integrals, stochastic properties, texture analysis

### ABSTRACT

This paper is concerned with the detection and separation of textures or stochastic information within an image. The texture is described mathematical by their stochastic. Because different kinds of stochastic processes generate the textures, they are represented by a system of coupled non-linear stochastic differential equations. If parameters a priori are not known, then their rank and derivations of this detect the stochastic. The distance of the interaction approximates the non-linearity. The entire knowledge does not exist and so the probability theory cannot be used, because the normalization implicit the knowledge of all possible properties. The usage of fuzzy measures compensates the loss of additivity. By fuzzy measures special properties without conditioned relationships to other properties can be described. The fuzzy measure allows the representation of different kinds of properties. The fusion of different kinds of properties related to their importance is realized by the fuzzy integral. The fuzzy integral is used as a new fuzzy measure and by iteration selected properties are isolated from the image successively. This method is applied for different stochastic or textural features within images such as clouds, contrails, error structures of surfaces of tissues, metals or medical images with stochastic structures.

### KURZFASSUNG

Der Vortrag beschäftigt sich mit dem Auffinden und der Separation von textuellen oder stochastischen Informationen in einem Bild. Die Textur wird mathematisch durch ihre Stochastik beschrieben. Da verschiedene Arten von stochastischen Prozessen die Textur erzeugen, werden diese durch ein System von gekoppelten stochastischen nichtlinearen Differentialgleichungen beschrieben. Wenn a priori keine Kenntnis über stochastische Parameter existieren, wird über die Bildung des Ranges die relevante stochastische Information aus dem Bild herausgezogen. Die Entfernung der Wechselwirkung wird über die Art der Nichtlinearität repräsentiert. Da eine vollständige Kenntnis über das System nicht existiert, kann mit Wahrscheinlichkeiten nicht gearbeitet werden, da die Normierung eine vollständige Kenntnis von allen Eigenschaften erfordert. Die Benutzung des Fuzzy-Maßes kann diesen Verlust der Additivität kompensieren. Mit dem Fuzzy Maß können verschiedenartige Eigenschaften beschrieben werden. Die Fusion dieser verschiedenen Eigenschaften erfolgt mit dem Fuzzy Integral. Das Fuzzy Integral wird als neue Fuzzy Maß benutzt und damit kann iterativ eine stochastische Eigenschaft aus dem Bild isoliert werden. Die Ergebnisse werden angewandt auf Erkennung von Texturmuster wie Wolken, Kondensstreifen, Störungen in Oberflächen von Stoffen, Metallen oder in medizinischen Bildern mit stochastischen Strukturen.

### 1 INTRODUCTION

Often exist information within an image, which are difficult to describe by edges, skeletons or other features. In this case a large part of the information contained in an image lies in their textural or stochastic structure. For the detection of such stochastic regions the information will be described by stochastic parameters such as variance, rank ordering, higher moments, martingales and Markov chains. The state-space model based on local values and the estimation theory for non-linear functions is the basis for such description. By such description it is possible to decompose the textures in the images in a sum of different described elementary stochastic properties. These different kinds of stochastic properties have to be represented by the same mathematical term. This is possible with the fuzzy measure, which maps the properties upon a measured unit of the same mathematical form. By the fuzzy integral selected parts of the set of stochastic properties are combined for description of new properties. Because the result of the fuzzy integral can be used as a new fuzzy measure, an iterative procedure can be constructed, to isolate regions with selected properties. This method is applied on images for detection and isolation of different cloud types, disturbances on surfaces of materials or for the analysis of medical images.

**2 REPRESENTATION OF THE STOCHASTIC INFORMATION BY KNOWN STRUCTURE**

The stochastic information in the image is represented in the state-space representation, where the local distance is used as the variable. There are two fundamentals for the description of the stochastics. Firstly it can be assumed, that interactions exist between different points in different distances. The interaction can be occur between nearer or further points (long distance and small distance interacting effects). On the other side such interaction can be of small-scale structure or large-scale structure. These effects occur, if relationships between different gray values exist based on interactions in the contents, such as clouds and their shadows. The relationships between different points in the image are considered in different distances, related to interactions, which may be far-reaching or shorter. Secondly it can be assumed that regions exist with especially stochastic properties. These regions are determined by an extension of special stochastic contents in an image. Special relationships between the points in the region do not exist. Such regions are e. g. the tree tops in a wood or the surface with waves on the sea. The scale respective the distance of the interaction is described by different kinds of non-linearity. For larger distances are used non-linear functions such as  $\log(x)$  or  $\sqrt{x}$ . For effects acting in shorter distances non-linear functions such as  $\exp(x)$  or  $x^n$  with  $n > 2$  are used. Here  $x$  is the distance of two points in an image given by the row index  $i$  and the column index  $j$ . In the second case, where only the region with stochastic properties exist the composed stochastic can be described by a sum of  $m$  coupled stochastic processes. For simplification, it can be assumed that the stochastic within this region can be derived from one colored noise generation process  $n(x)$ . The system of coupled stochastic non-linear differential equations for a selected region with stochastic contributions is of the form:

$$\begin{aligned} dy_1(x) &= \mathbf{F}_1(y_1(x), y_2(x), \dots, y_k(x)) dx + G_1 n(x) \\ dy_2(x) &= \mathbf{F}_2(y_1(x), y_2(x), \dots, y_k(x)) dx + G_2 n(x) \\ &\dots\dots\dots \\ dy_m(x) &= \mathbf{F}_m(y_1(x), y_2(x), \dots, y_k(x)) dx + G_m n(x) \end{aligned}$$

Here,  $y_1, y_2, \dots, y_m$  are the  $m$  components of the stochastic and  $\mathbf{F}_1, \dots, \mathbf{F}_m$  are the  $m$  non-linear matrices for the relationships between the components. The  $m$  coefficients  $G_1, \dots, G_m$  are the gain factors for the noise  $n$  and  $x$  represents the pixel-point or the region in the image. If interaction exist between different points as in the first case described above, the coupled system of equations are modified by

$$\begin{aligned} dy_1(x) &= \mathbf{F}_1(y_1(x), y_2(x), \dots, y_r(x)) dx && r < k \\ dy_2(x) &= \mathbf{F}_2(y_1(x), y_2(x), \dots, y_r(x)) dx && r < k \\ &\dots\dots\dots \\ dy_s(x) &= \mathbf{F}_s(y_1(x), y_2(x), \dots, y_r(x)) dx && r < k \\ dy_{s+1}(x) &= \mathbf{F}_{s+1}(y_1(x), y_2(x), \dots, y_n(x)) dx + G_{s+1} n(x) \\ &\dots\dots\dots \\ dy_m(x) &= \mathbf{F}_m(y_1(x), y_2(x), \dots, y_n(x)) dx + G_m n(x) \end{aligned}$$

Besides this system for the data exists a stochastic equation for the acquisition of the data. In vector representation the system of stochastic differential equations can be written in a stochastic integral representation of digitized form by:

$$\mathbf{y}(x_{l+1}) - \mathbf{y}(x_l) = \left[ \mathbf{F}(\mathbf{y}(x_l)) + \int_{x_l}^{x_{l+1}} G_l(x) dW(x) \right] (x_{l+1} - x_l)$$

$x_{l+1}$  and  $x_l$  are neighboring points, also e. g.  $x_{i,j+1} - x_{i,j}$ ,  $x_{i,j} - x_{i,j-1}$ ,  $x_{i+1,j} - x_{i,j}$ ,  $x_{i-1,j} - x_{i,j}$ ,  $x_{i+1,j+1} - x_{i,j}$ ,  $x_{i-1,j-1} - x_{i,j}$ ,  $x_{i-1,j+1} - x_{i,j}$  and  $x_{i+1,j-1} - x_{i,j}$ .

The information within the pixel  $\mathbf{y}(x_{l+1})$  is obtained by the observation process  $z(x_k)$ . The digitized observation values  $z(x)$  including disturbances in the CCD matrix and the AD converter are given by:

$$z(x_{k+1}) = \int_{x_k}^{x_{k+1}} \mathbf{H}(x) \mathbf{y}(x) dx + \int_{x_k}^{x_{k+1}} C(x) dV(x)$$

W(x) and V(x) are independent Wiener processes. The non-linearities  $\mathbf{H}(x)$  are determined by a priori knowledge. For the solution of the system of non-linear stochastic differential equations the martingale representation [2, 3] is applied. By this method the effect of approximation is clearly to understand because the calculation remains in the same space as in the model given. This is not the case for most other methods such as by solving the Fokker–Planck equation. In [2] it has been derived, that for a square integrable martingale the process  $p_x(z, y)$  can be written in an integral representation of the form

$$p_x(z, y) = p_0 + \int_0^x P_s ds + n_x$$

where  $N = (n_x, \mathfrak{R}_x)$  is a martingale measured by sigma algebra  $\mathfrak{R}_x$  and  $P_s$  is a stochastic process. Because the stochastic differential equations depend only on z, we obtain for the expectation value of  $p_x(z, y)$

$$\langle p_x(z, y) \rangle = \mathbf{E} [ p_x(z_x) | \mathfrak{R}_x^y ] .$$

By using the martingale method the effect of approximation is clearly to understand because for the calculation the same space is used as for the model.

### 3 REPRESENTATION OF THE STOCHASTIC INFORMATION BY UNKNOWN STRUCTURE

Often parametric information about the stochastic within the image does not exist. In such a case the stochastic structure can be described by using non-parametric methods [4]. An efficient method for this is the rank description [5]. For images with gray values  $x_{i,j}$  of the pixel points  $(i, j)$  the rank  $R_{k,l}$  is given by

$$R_{i,j}(M, N) = \sum_{l=-M}^M \sum_{k=-N}^N u(x_{i,j} - x_{l,k}) \quad \text{with the step function} \quad u(z) = \begin{cases} 0 & \text{for } z < 0 \\ 1 & \text{for } z \geq 0 \end{cases}$$

Here the selected points surrounding the pixel point  $x_{i,j}$  are  $k = \{i-M, \dots, i-1, i+1, \dots, i+M\}$  and  $l = \{j-N, \dots, j-1, j+1, \dots, j+N\}$ . The rank  $R_{i,j}$  shows how many of the pixels in the region  $2N \times 2M$  have a gray value less than the selected value  $x_{i,j}$  in this region. The distributions of all rank values are obtained by shifting  $x_{i,j}$  over the entire image. A particular stochastic component  $S_{i,j}(g)$  with the threshold g is obtained if only those rank values  $R_{i,j}$  are selected which validate  $x_{i,j} > g$ . Different kind of stochastic are characterized by changing the threshold g and gives:

$$S_{i,j}(g, M, N) = \sum_{l=-M}^M \sum_{k=-N}^N u(|x_{i,j} - g| - |x_{l,k} - g|) u(x_{i,j} - x_{k,l})$$

Because  $u(x) = \frac{1}{2} \operatorname{sgn}(x) + \frac{1}{2}$  and  $u(x_{i,j} + x_{k,l} - 2g) = 1$ , this equation is fulfilled if and only if  $|x_{i,j} - g| \geq |x_{k,l} - g|$  and  $x_{i,j} > x_{k,l}$  or  $|x_{k,l} - g| \geq |x_{i,j} - g|$  and  $x_{k,l} > x_{i,j}$ . So for the number of the neighboring values larger than  $x_{i,j}$  and a threshold g can be written:

$$S_{i,j}(g, M, N) = \sum_{k=-N}^N \sum_{l=-M}^M u(x_{i,j} + x_{k,l} - 2g)$$

$S_{i,j}(g)$  is used as a new value for the pixel point  $(i,j)$  and a new kind of image is produced, where the fluctuations of the desired stochastic component have been reduced. Based on rank representation four types of stochastic can be separated:

- $S(g)$  is calculated for different values of g

- The difference is calculated for different levels  $S(g_m) - S(g_n)$
- A repeated calculation of rank by using  $S_{i,j}(g, M, N)$  deliver a new value for  $(i,j)$
- The variance of  $S_{i,j}(g, M, N)$  is calculated over the direct neighborhood and used as a new value for  $(i,j)$ .

The stochastic properties obtained by parametric or non-parametric methods can be used together for the fusion.

#### 4 MAPPING OF THE STOCHASTIC INFORMATION ON A FUZZY MEASURE AND FUZZY FUNCTION

The description of all components of the stochastic is very difficult and mostly impossible, because they are not fully known. This incomplete knowledge basis implicates that different parts of information have different importance. This can be described mathematically by the fuzzy measure. On the other hand, for the fusion of different kinds of stochastic a representation of stochastic properties in the same expression is necessary. This is also possible by the fuzzy measure. For a mathematical basis of the fuzzy measure will be used the Lebesgue measure. The stochastic properties with finite values can be measured by the Lebesgue measure. These properties are measurement on the open interval  $b_v - a_v$ , related to the pixel distance. This is related to the elementary geometrical content  $|i|$  of

$$|i| = \prod_{v=1}^n (b_v - a_v) = \sum_{v=1}^n (-1)^v c_1 \cdots c_n$$

Here are  $c_v$  the interval (pixel) points and the Lebesgue measure is the summation of the contents of these intervals. The characteristic function describes the membership for a measure  $i$  and is 1 if  $i$  is a member of the set and 0 if not.

The characteristic function for the set  $\bigcap_{k=1}^n i_k$  is given by  $\prod_{k=1}^n i_k(\xi) = \inf_{1 \leq k \leq n} \{g(\xi_k)\}$  with  $\xi$  as a property represented by the

membership of a set for a selected condition. This measure gives the possibility to define a function of arbitrary pieces of intervals as desired but it is an additive measure. This contradicts the inclusion of the importance for the decision. Therefore, this measure has to be extended. A coupling between the components represents the importance of a property. Such a coupling by a factor  $\lambda$  is given by the fuzzy measure (densities)  $g(\xi)$  introduced by Sugeno[6] as

$$g_\lambda(\xi_1 \cup \xi_2) = g_\lambda(\xi_1) + g_\lambda(\xi_2) + \lambda g_\lambda(\xi_1)g_\lambda(\xi_2)$$

By the fuzzy measure the properties described by different kinds of relationships are mapped into the closed interval  $[0,1]$ . The different types of stochastic are collected in a set of properties, where the members of this set are fuzzy in their contribution for the determination of a selected area. For an optimized decision all relevant stochastic properties have to be fused. For this fusion stochastic properties are not only described by a fuzzy measure but also by a fuzzy function. If the fuzzy property is more related to a region, then a fuzzy measure is used; if a stochastic property is better described by a particular distribution of gray values, then this is represented by a fuzzy function. The factor  $\lambda$  has an effect similar to a weight factor. If  $\lambda = 0$  then the fuzzy measure is equal to the probability measure.  $\lambda$  calculated by  $\lambda = (1 + \lambda g(\xi_1))(1 + \lambda g(\xi_2)) - 1$  is used as a substitution for the loss of additivity. For a set of elements  $A = \{\xi_i\}$  the relationship above can be used recursively and gives [6]:

$$g(A) = \sum_{i=1}^n g(\xi_i) + \lambda \sum_{i=1}^{n-1} \sum_{j=i+1}^n g(\xi_i)g(\xi_j) + \cdots + \lambda^{n-1} g(\xi_1) \cdots g(\xi_n)$$

This can be written as product

$$g(A) = \frac{1}{\lambda} \left[ \prod_{\xi_j \in A} (1 + \lambda g(\xi_j)) - 1 \right] \quad \text{where } \lambda \neq 0 \quad \text{and} \quad 1 + \lambda = \left[ \prod_{\xi_j \in A} (1 + \lambda g(\xi_j)) \right]$$

The non-linear equation for  $\lambda$  can be solved iterative by coupling two properties  $\xi_i$  and  $\xi_m$  and using the result for the next coupling. Properties especially capable for the representation by a fuzzy measure are stochastic properties such as:

- logical functions of some bitmaps of data within an interval,
- estimated values of non-linear filtering higher/lower as a threshold
- regions given by retransform of a wavelet representation with adapted coefficients
- stochastic values obtained by the subtraction of the original and estimated value

The fuzzy functions are constructed as fuzzy values over a region. Normally these fuzzy functions are described by a characterisation over a threshold. Outside of such a characteristic threshold the values depend very weakly on the real value. Inside the interval the values generate fuzzy properties for the selected condition. For the fuzzy function properties are used such as:

- differences of data values in different distances
- stochastic changes of the data values in related areas
- differences of the stochastic in different directions
- number of ranks related to different distances
- values obtained by wavelet transformation of a selected parameter

These fuzzy functions are also normalized and mapped on an interval [0,1]. Whereas the fuzzy measure is better adapted for effects represented in special regions, the fuzzy function characterizes the stochastic change over a region of a fuzzy measure. If the property is more functional than it is represented by a fuzzy function  $h(\xi_{k,l})$  over a region of a fuzzy measure  $g(\xi_{k,l})$ . The values are combined in the similar manner as with the fuzzy measure. For clarifying the usage of fuzzy measures and their related fuzzy function two examples may be used. If the image values are divided in bitmaps, the two highest can be used for the fuzzy measure to describe a region and the two lowest bitmaps can be used as the fuzzy function describing the stochastic in the region. The connection of both describes selected stochastic properties as e. g. the detection of smoke of a fire. Another case is the description of a combination of histogram values over very small regions obtained with different intervals by the fuzzy measure, which may be combined with the region obtained by the assumption of a limited variance as the fuzzy measure.

### 5 FUSION OF DIFFERENT KINDS OF STOCHASTIC INFORMATION BY FUZZY INTEGRAL

The fusion of stochastic information is based on elementary stochastic information if possible. Mostly the stochastic information exists only in a composed form and has to be decomposed. This decomposition is also possible with the help of the fuzzy integral. In the fuzzy integral a part of the fuzzy measure and a part of the fuzzy function are linked together whereas other properties in such a combination will be suppressed. By the fuzzy integral values over the possible region of a selected stochastic, represented by a fuzzy measure  $g(\xi_{k,l})$ , are connected with the stochastic values  $h(\xi_{k,l})$  defined over this region. Therefore, the fuzzy integral represents a functional relationship between the fuzzy measure and the fuzzy function. For the fuzzy integral the old definition of the fuzzy integral of Sugeno [6] will be used, because it is well adapted to the problem of detection of stochastic properties. With Sugeno [6] the fuzzy measure is combined with the fuzzy function in the form (written as a stylised  $f$ ):

$$\int_A h_\alpha(\xi) \oplus dg = \sup_{\alpha \in [0,1]} \{ \min [\alpha, g(A \cap H_\alpha)] \}, \quad H_\alpha = \{ \xi | h_\alpha \geq \alpha \}$$

Here,  $h_\alpha$  is the cut of  $h$  at the constant  $\alpha$ . For  $h_\alpha(x)$  the values over the regions of selected pixel-points are used, representing a stochastic property.  $\alpha$  is the threshold where the assumption is fulfilled, that the property is used in the minimal condition. The region  $A$  is given as the data region where surely a specific stochastic property is expected. It may also be the entire image for the pixel region and the entire possible range for a fuzzy function.

The important property of a fuzzy measure and a fuzzy function is that its different kinds of value are mapped upon the closed interval [0,1]. The value obtained by the fuzzy integral also is defined in the interval [0,1]. This gives the possibility to use the result of a fuzzy integral as a new fuzzy measure  $g_2$

$$g_{2_{k,l}}(A) = \int_A h_\alpha(\xi) \oplus dg_1$$

This newly produced fuzzy measure describes a region, which can be connected with another stochastic property represented as a fuzzy function. In a similar way the result of the integral can also used as a fuzzy function. Consequently, a new fuzzy function  $h_{\alpha_2}$  can be constructed from the fuzzy integral:

$$h_{\alpha_2} = \int_A h'_{\alpha'}(x) \oplus dg$$

In the following step the next region of another property is combined with this fuzzy function  $f_2$ . In such a way a set of fuzzy functions  $\{h_{\alpha_1}, h_{\alpha_2}, \dots, h_{\alpha_m}\}$  can be obtained by fuzzy integrals such as a set of fuzzy measures  $\{g_1, g_2, \dots, g_m\}$ . The components of both sets are combined for selecting specific stochastic properties. That way, coincidences of the properties contained in the fuzzy measure and the fuzzy function are filtered with the fuzzy integral.

The summation of all combinations of fuzzy measures with fuzzy functions makes sure that all possible properties in all combinations, which should be considered, are extracted. In such a way an image is obtained, where the (gray) values represent a measure of the membership to the stochastic. With the help of a threshold it can be decided, which pixel of the data belongs to a selected stochastic property.

For the understanding of the usage of the fuzzy integral an example may be considered. Roughly spoken, the fuzzy integral selects the largest value of all minimal fuzzy values mapped upon the fuzzy measure restricted by a defined area and cut by the constant  $\alpha$ . Thus the importance of a property is included. This includes the neglecting of contributions if a value is less than an assumed threshold. The region may be described by the difference of two images taken at different time or in different frequency regions succeeding a given threshold. This region describes a changing area during time or area distinguished by different frequencies related to given parameters. This area is connected with stochastic values, produced by a combination of mathematical moments as variances or skewness combined with rank or sign statistic, representing selected stochastic properties. The obtained result is then used as a new region and this procedure is continued until the region will no more essentially changed.

## 6 APPLICATIONS AND RESULTS

For the detection of different types of clouds MOS, AVHRR and ATSR images are used. Upper figures show on ATSR images, made available by the head of CLOUDMAP, Prof. Jan-Peter Muller, UCL, London, as with the help of the method given above, by estimation of the stochastic structures several regions of the image can be distinguished. The combined

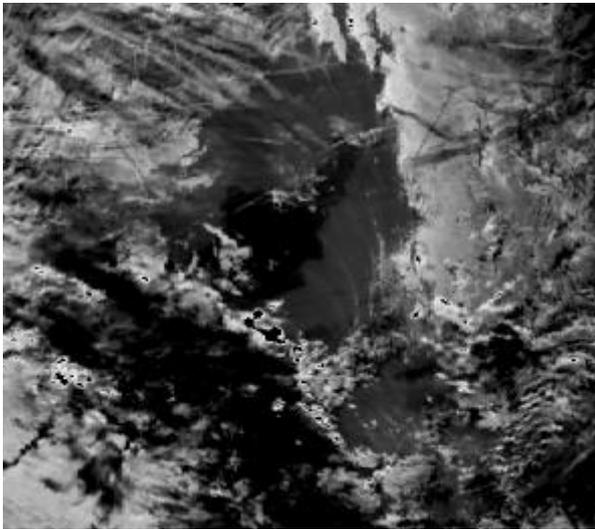


Fig. 1 Image of differences of two spectral channels

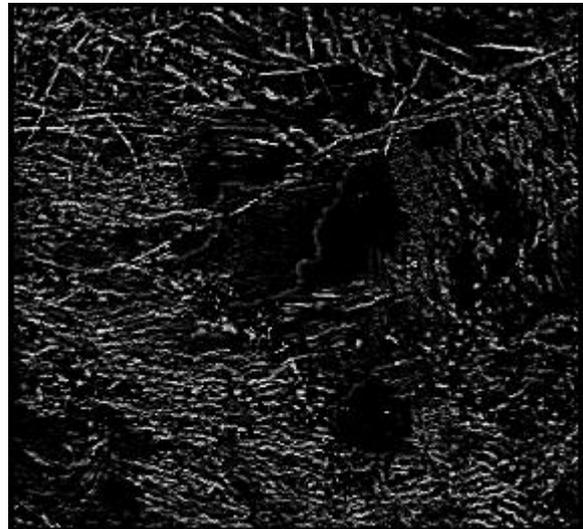


Fig. 2 Extraction of lines structures

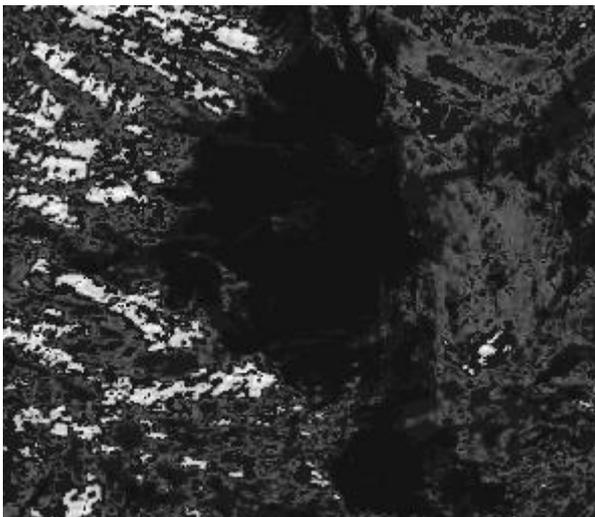


Fig. 3 Extraction of clouds with stripe structures

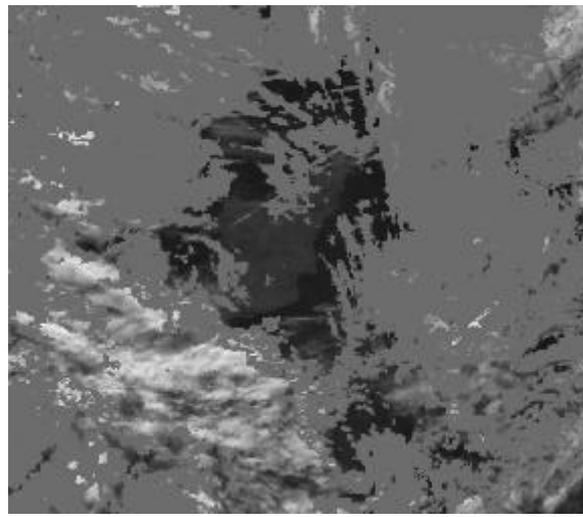


Fig. 4 Extraction of cumulus like cloud structures

structural effects are the basis to get different regions of structures. By an iterated applied procedure for detection small structures the contrails would be isolated. It is also important for this to eliminate cloud with stripe structure as shown in Fig. 3. The small structure within clouds, produced by wind effects, have also to eliminate, as shown in Fig. 4. This method is also applied for the detection cirrus clouds and the estimation of the boundary of such a thin cloud with normally not visible edges.

Another field of application is the image of treetops. Health and age of the forest are correlated with the stochastic properties within images of the treetops taken from towers. The stochastic properties of treetops are separated by non-linear filtering and non-parametric methods connected with non-stochastic properties such as brightness differences produced by different kinds of trees or forest aisles. The separated textural properties are composed for new stochastic properties describing the forest status or damages of treetops by air pollution. Different kinds of stochastic properties are mapped on the closed interval [0,1] and represented by fuzzy measures or related fuzzy functions. Such properties are: differences of values of neighbouring pixel, differences of the stochastic in different directions and the rank of grey values. By the fuzzy integral the so represented properties are fused with regard to their importance. The fuzzy integral fulfils the mathematical properties of a fuzzy measure. Therefore, by this fusion new fuzzy measures are generated. By an iterative process, stochastic properties are separated and then selected components combined for a complex stochastic property. Changes of distances between elementary structures by different observation angles are corrected.



Fig. 5 Original image of forest of different trees variably old



Fig. 6 Pines are selected from other trees and by age



Fig. 7 Selected birch trees (weed of forest) from other trees



Fig. 8 Selected young pines in a forest aisle

The so obtained results are represented by bitmaps representing special structures of the forest such as birches in pine forest or forest areas separated by age or illness of treetops. The method is applied for pines and mixed forests.

## 7. CONCLUSIONS

In the paper the stochastic properties are applied to select regions within an image. The method has another fundament as the usual methods of digital image processing as e.g. given in [7]. The stochastic properties are described by a system of stochastic differential equation, which are solved by martingale technique. With the help of the fuzzy measure and the fuzzy function different kinds of stochastic properties are represented by the same term. The fusion is achieved by the fuzzy integral in the form of an extended Lebesgue integral. Because the fuzzy integral has mathematical properties as the fuzzy measure, an iterated process is constructed to isolate selected stochastic properties within an image. The method is applied for detection of different cloud types, disturbances in surfaces, detection of smoke and estimation of the state of the forest by the treetops.

## ACKNOWLEDGEMENTS

I would like to thank the EU for the support (EU Contract no. ENV4 CT97-0399) which was the basis for this investigations.

## REFERENCES

- [1] Basseville, M., Nikifirov, I.V., 1993. Detection of Abrupt Changes: Theory and Application, P T R Prentice Hall, Englewood Cliffs, New Jersey
- [2] Liptser ,R. Sh. and Shirayayev A. N.: "Statistics of random processes", I-II, Springer-Verlag , (translated from the Russian) ,1977-78.  
Protter, P. "Stochastic Integration and Differential Equations", Springer-Verlag, 1995
- [3] Hetzheim, H. 1993, "Using Martingale Representation to Adapt Models for Non-Linear Filtering", Proceedings of ICSP'93, International Academic Publ. , Beijing , pp.32-35, (1993).
- [4] Bickel, P. J ., Klaassen, C. A. J., Ritov, Y. and Wellner, J. A. " Efficient and Adaptive Estimation for Semiparametric Models",The Hopkins Univ.Press, Baltimore and London, 1993.
- [5] Hetzheim, H. , Dooley, L. S.Structured non-linear algorithms for image processing based upon non-parametric descriptions, Conference Proceedings of Second Asian Conference on Computer Vision(ACVV'95)
- [6] Sugeno, M. Theory of fuzzy integrals and its application, Doctoral Thesis, Tokyo Institute of Technology, 1974.
- [7] Teubner,J. "Digital Image Processing", Prentice Hall, New York, 1993