EVENTS-BASED IMAGE ANALYSIS FOR MACHINE VISION AND DIGITAL PHOTOGRAMMETRY

Yury V. Visilter, Sergei Yu. Zheltov, Alexander A. Stepanov

(State Research Institute of Aviation Systems, Moscow, Russia)

KEY WORDS: Image Analysis, Object Recognition, Bayesian theorem.

ABSTRACT:

A new approach for model-based image analysis called the *Events-based image Analysis* (EA) is proposed. From EA point of view, any certain procedure of image understanding can be interpreted as a procedure of *evidence fusion*. Any fact about the whole image, about its part or even about one proper pixel can be the *evidence*, and the any proposition about the scene observed is the *hypothesis* that to be tested based on these evidences. In this paper the EA formalism was outlined in the Bayesian terms. This approach allows to compose the power of sample-based methods and the flexibility of model-based methods without the direct comparison of objects or images. The most important properties of EA procedures are the following: usage of generic models, usage of hierarchical models and easy fusion of non-homogeneous information. Based on EA ideas the complex technique for house detection was proposed. It provides the easy fusion of contour and intensity information for 3D-model validation.

1. INTRODUCTION

Any certain engineering technique can be considered as a combination of "*method*" and "*state of art*". Here, "method" is a regular part of technique that to be the best for a common class of problems involved. In the contrary, the term "state of art" means the part of technique that reflects the peculiarity of the concrete problem and due to this - the skill of the developer. It is very attractive to define the common method for the most wide spectrum of problems, to prove its' optimality and then to concentrate on the state of art only.

Some object models (of different types) are developed for detection and measurement of artificial objects, and the proper measurement presumes the estimation of numeric parameters of these models. So, the detection methods based on such models to be preferred. However, though some of model-based detectors are developed, usually they contain the heuristic matching procedures without any of optimal assumptions. Because of this reason, most powerful and robust object detection techniques do match the sample images but not the models.

In the earlier works our group intensively used two well-known matching techniques: the *Pytiev Morphology* and the *Hough Transform*. The Pytiev Morphology provides the most invariant detection of objects by their samples, but it can not work with models. The Hough transform (HT) and the Generalized Hough Transform (GHT) support the efficient model-based contour analysis, but can not use any other type of information contained in images. The common "method" called the "*Events-based image Analysis*" (EA) was developed to compose the Pytiev's optimal state of problem and the methodology of Hough transform. EA is a "method" for the most generic model-based image analysis while its' "state of art" is connected with a choice of the adequate models.

From EA point of view, any certain procedure of image understanding can be interpreted as a procedure of *evidence fusion*. Any fact about the whole image, about its part or even about one proper pixel can be the *evidence*, and the any proposition about the scene observed is the *hypothesis* that to be proved or escaped based on these evidences. There are many possible ways to provide the such fusion. However, the Bayesian approach is the most popular and meaningful. So, we shall outline the EA formalism in the Bayesian terms.

In section 2 we outline the basic idea of Bayesian EA. In section 3 the EA-approach for house extraction is described as an example of EA-application for photogrammetry and machine vision.

2. EVENTS-BASED IMAGE ANALYSIS: DETECTION WITHOUT COMPARISON.

From the most generic point of view, the set of possible object recognition procedures can be separated into two principal groups: methods that use the object comparison and methods that do not use it.

This distinction is easy to show for the case of the feature-based object classification. Methods of closest neighbors based on the comparison of the each new vector that characterizes the object with some sample vectors that characterize known object classes. A corresponding distance or a closeness measure is computed here to be a criterion of classification. So, the reliability of object recognition (detection) is determined by the comparison of the objects in some metric space. On the other hand, the statistical classificators immediately use the probability distribution functions to estimate the reliability of object classification. Some set of samples can be used at the training stage but at the stage of decision making any comparison with samples is not of use.

The most important limitation of the comparison-based methods is that we can compare only the objects of the same type (two images, two contours, two vectors, two wire models and so on). So, we can not compare the image and the model.

In our opinion, the most powerful comparison-based detection technique is the Pytiev's morphological analysis (Stepanov at al., 1994) that really provides the invariantness of object detection. Due to this it demonstrates the possibilities and disadvantages of comparison-based methods with great expression. The main idea of Pytiev's morphology is the following. Let the images will be the elements of some Hilbert space $IM \sim L^2$. So, one can speak about an image *norm* $||Im||$ and a distance between the images $||Im1-Im2||$. Let also some convex and closed image set $Z \in IM$ is given. Then for any image $Im \in IM$ there is the unique image $Im' \in Z$ such that $||Im'-Im||=\min\{||Im''-Im||, Im'' \in Z\}$. It is easy to see that this mapping $v(Im): Im \to Z$ is a *projecting operator* in the (algebraical) sense that $v(v(Im))=v(Im)$. So, we can note $Im'=Pr_Z(Im)$, i.e. Im' is the projection of Im onto the Z. Using the image projection notion some special closeness measure $K(Im,Z)$ (*the morphological correlation coefficient*) can be defined. It is analogous to the usual correlation:

$$K(Im,Z)=||\inf(Im,Pr_Z(Im))||/||\sup(Im,Pr_Z(Im))||$$

and has the following useful properties: 1) $0 \geq K(Im,Z) \geq 1$, $Im \in IM$, $Z \in IM$; 2) $(K(Im,Z)=1) \Leftrightarrow (Im \in Z)$.

The basic advantages of the morphological correlation coefficient are connected with the possible full account of the registration conditions. Let the registration model is described by some transform $s \in S$ where S is a semigroup of transforms and the object model is $M=\{Im^M\}$ (object is described by its' sample). The *Pytiev's morphological shape* of any image Im will be $Z_{Im}=\{Im'=s(Im), s \in S\}$. So, the morphological correlation coefficient $K_S(Im',Im)=K(Im',Z_{Im})$ provides the correct comparison between any test image Im' and the given sample $Im=Im^M$ under the condition of transformation $s \in S$.

For instance, let consider the generic model of radiometric distortions. In the formal way any image is a 2D-function of intensity distribution and can be represented as $f(x,y)=\Sigma(a_i \times \chi_i(x,y))$, where χ_i is an *indicator* of the i-th region of the cadre tesselation ànd a_i is a color (intensity value) of this region. So, the set of images "of the same shape" has a form:

$$Z=\{f''(x,y)=\Sigma(b_i \times \chi_i(x,y)), \forall\{b_i\}\}.$$

Then the projecting transform is a parametric one and has the form: $b_i=b(a_i)$, where i=0..C-1; C is a number of colors in the image. For any image g(x,y) the projection $Pr_f(g)$ is defined by the parameter vector $\mathbf{b}_g$:

$$b_i=(\iint(g(x,y)\chi_i(x,y))dxdy)/(\iint\chi_i(x,y)dxdy),\ i=1..C-1,$$

that is easy to compute. Since the parameters of projection are computed the morphological correlation coefficient K(g,f) is computed immediately.

So, we see that the Pytiev's morphology decides the problem of the invariant object detection in the case of object model $M=\{Im^M\}$ under the regular registration model S. However, when the model M does not satisfy any special conditions, computation of the $Pr_M(Im)$ is too hard because we need to compare the test image Im with the each element Im' from M to find the closest one.

Consider this problem applying to non-comparison-based techniques. The simple example is a classic pattern analysis using Hough Transform (Houle and Malowany, 1989).

Let the set of images to be analyzed is a set of planar dot patterns and it is required to detect all straight patterns in it. The *straight pattern* here means the sub pattern that contains a number of points that lie on the same straight line. It is very fuzzy and flexible model M because neither the number nor the location of points on the line is not defined. So, we can not use any sample here. The registration model allows the affin transforms of the image plane.

The Hough Transform (HT) is a well-known technique for object detection in the parameter space. It uses the parameter space (ρ,θ) of the normal line equation $X\cos(\theta)+Y\sin(\theta)=\rho$. The set of parameters (ρ,θ) of all possible lines that intersect in some proper point (x,y) of the image plane corresponds to a sinusoidal figure in the space (ρ,θ). This figure is called the *spread function*.

The idea of HT is to accumulate the votes of the pattern points in the parameter space through the simple summation of their spreads. If two points of the pattern belong to some line (ρ_i,θ_i) then their spreads intersect in the point (ρ_i,θ_i) in the Hough space (ρ,θ). So, the value of the resultant *accumulator function* $A(\rho,\theta)$ in the each point (ρ_i,θ_i) is equal to the number of points of the pattern that lie on this line (ρ_i,θ_i). Thus, if the pattern contains *m* straight patterns, it will be *m* local maxima in the Hough space.

It is very efficient technique that provides the invariant detection of the straight patterns without any comparison with samples. The Hough Transform does not require any sample Im^M because it immediately accumulates the votes for a model M. So, *techniques that do not use the comparison can work directly with generic models of objects.*

The Events-based image Analysis (EA) approach was developed to generalize this important property of Hough transform for a common case of object detection. The essence of EA is the following.

Let we have some image Im, and it is required to determine a posterior probability of some hypothesis H about the scene observed. Then the Bayesian formula takes the form:

$$P(H/Im)=[P(H)\times P(Im/H)]/[P(H)\times P(Im/H)+P(H^c))\times P(Im/H^c)], \qquad (1)$$

where H^c means "not H".

Image Im is also considered (in the spirit of Probability Theory) as an *event*, or, in other words, we consider the event E(Im) that is connected with this image Im. This event E(Im) consists of some different events occurred in the process of low-level image analysis.

While the any essential fact derived from image analysis is the *event* e_k, the event E(Im) will be the intersection

$$E(Im)=e_1\cap e_2\cap..\cap e_K, \qquad (2)$$

where K is the total number of such events. So, we need only (1) and (2) to test any hypothesis H about the image Im.

If one supposes that events $\{e_k\}$ are independent in general then (1) and (2) supply

$$P(H/\mathrm{Im})=\frac{P(H)\times\prod\{P(e_k/H)\}}{P(H)\times\prod\{P(e_k/H)\}+P(H^c)\times\prod\{P(e_k/H^c)\}},\ldots(3)$$

where $\prod\{x_k\}=x_1\times x_2\times\ldots\times x_K$.

From this point of view the Hough Transform, Generalized Hough Transform, Serra morphology, Pytiev's morphology and many other popular techniques are the Bayesian EA-procedures that differ in events analyzed, hypothesis tested and probability models used.

The most important properties of EA procedures that are principally improper for the comparison-based techniques are the following:

- the usage of generic models;
- the usage of hierarchical models;
- the usage of non-homogeneous information.

The first one is provided through the accumulation of evidences immediately for the model-based hypothesis. The most important result here is that the assumption of event's probability independence in general is enough to provide the possibility of parallel independent accumulation of events' evidences.

The usage of hierarchical models based on a hierarchical application of Bayesian formula.

The usage of non-homogeneous information is clear enough but usually connected with a coarsening of the model of real situation. The non-homogeneous image data means a set of data from different physical image sources or/and from different image processors. Let we have N channels of registration and L levels of data abstraction. *Level of data abstraction* is a form of information representation (image, contour preparation, dot pattern, parameter space, feature vector, etc.). Let the complex *model of object* is described as a set of propositions $\mathbf{M}=\{M_l^i\}$, i=0..N; j=1..C, l=0..L, that the object must satisfy to. The notation M_l^i means that this proposition takes place in i-th channel at l-th level of abstraction if the object of model **M** is observed.

Then a posterior probability (3) takes a form:

$$P(H / IM) = \{P(H) \times \prod_{l=1}^{N} \prod_{l=1}^{L} P(M_l^l / H)\} /$$
$$/\{P(H) \times \prod_{l=1}^{N} \prod_{l=1}^{L} P(M_l^l / H) + P(H^C) \times \prod_{l=1}^{N} \prod_{l=1}^{L} P(M_l^l / H^C)\}, \cdots (4)$$

So, the events-based image analysis provides a generic framework for non-homogeneous information analysis.

3. 3D-MODEL TO IMAGE MATCHING FOR HOUSE DETECTION.

The problem of automatic 3D-model to image matching is discussed in many papers and publications. Let consider two of them that present the most pure concepts of such matching. While one presumes that the complete 3D wire frame model of the house and full camera geometry are known, the "*prediction*" of 2D-contours of the house image can be build. Then it can be matched to the real contour preparation on the observed image. It is not a trivial task due to the weak correspondence between the ideal contours and the production of real edge detectors. Such sophisticated contour-based matching technique is described in (Huertas, Bejanin and Nevatia, 1995). Its robustness strongly depends on the quality of initial contour preparation. In the paper (Mueller and Olson, 1995), the intensity-based correlation approach is presented. In this way the 2D prediction is an intensity image and so one can reduce "model-to-image matching" problem to the well known "image-to-image matching" problem. However, to predict the intensity values on the model image authors had to make both the geometric and the radiometric prediction. The latter problem is sophisticated too because, even the 3D-model includes the plane surfaces only, it requires to estimate the color and the reflectivity of these planes as well as the sun luminance characteristics. The results seem to be satisfactory enough, but it is the rare case when the reflectivities of the model facets are precisely known.

As shown above, the Pytiev morphology provides a way for comparison of images by their "shape" but not immediately by pixel intensities. The "shape" of the intensity image is equivalent to 2D-area tessellation and can be described by a set of homogeneous areas that cover the image and pair-wise not intersected. The projection of 3D wire frame onto the image plane determines the unique Pytiev's "shape". Then the test patch of the real image must be "*projected*" (in the Pytiev sense) onto this shape. This "*morphological projection*" will be the required model approximation to be compared with the real image patch. Thus, the Pytiev technique allows to realize the intensity-based model-to-image matching using only the geometric prediction (without the any of radiometric knowledge).

Let consider the simple case of planar facets and Lambert's reflection model. It means that the intensity of reflected light is just proportional to the angle of the facet inclination and, consequently, the intensity of any image region corresponded to the facet must be constant. Under these assumptions, the morphological projection can be obtained in the most simple way, through the computation of the average values of image intensity over the each region of the "shape". As we understood, Mueller and Olson used the analogous technique (to compare with their approach) and found it unsatisfactory due to false detections occurred. These results are correct if the morphological projection is used as a prediction and compared with image by the usual correlation way. However, the real success of the intensity-based model-to-image matching takes place only if two following facts are proved:

1. The intensity over the each of facet 2D-projection (region) is homogeneous enough;
2. The edges between different facets are expressed enough.

Contour data and intensity data make up the non-homogeneous information set. So, they can be fused in the EA-manner as described before. To do this we need to agree some probabilistic model of object. Let the intensity of pixels on the each facet projection is described by a Gaussian distribution. Let the probabilities of contour point at the expected contour and out of the expected contour are known a priory (from expert analysis). We think that the assumption of independence of pixel events is an appropriate coarsening of reality. These assumptions lead the following algorithm of model-to-image matching:

1. Build the 2D-projection of the object's wire-frame onto the image plane to define the model image "shape". Project (in the Pytiev sense) the registered image onto the model "shape". Estimate the parameters of intensity distribution using the mid-level approximation as a set of average values.
2. Build the contour preparation of the image.
3. Evaluate the non-homogeneous criterion P(H/Im) (4) that characterizes the quality of model-to-image matching.

This approach has some advantages in comparison with the discussed predictive approach:

- A prior radiometric information is not required.

- The estimation of sun emission characteristics is not required.
- The desired contours are tested to improve the matching results.
- The shadows' evidences can be taken in account immediately in the matching process.

The last point is clear from the look at the "shape" of any typical image of a house with a shadow. We can account the shadow casting at the stage 1 of our matching algorithm (while creating the "shape"). Additionally, we can especially check that the shadow region of the mid-level approximation is dark enough.

The following improvement of this technique is connected with the account of non-pixel events. For example, one can extract the straight lines' segments on contours and consider them as contour events.

4. CONCLUSION

A new approach for model-based image analysis named the *Events-based image Analysis* (EA) is proposed. From EA point of view, any certain procedure of image understanding can be interpreted as a procedure of *evidence fusion*. Any fact about the whole image, about its part or even about one proper pixel can be the *evidence*, and the any proposition about the scene observed is the *hypothesis* that requires to be proved or escaped based on these evidences. In this paper the EA formalism was outlined in the Bayesian terms.

This approach allows to compose the power of sample-based methods and the flexibility of model-based methods without the direct comparison of objects or images. The most important properties of EA procedures that are principally improper for the comparison-based techniques are the following:

- the usage of generic models;
- the usage of hierarchical models;
- the usage of non-homogeneous information.

Based on EA ideas the complex technique for house detection was proposed. It provides the easy fusion of contour and intensity information for 3D-model validation. This technique was realized as a low cost application on IBM PC and preliminary tested using a set of images of Ufa city (Russia). The preliminary results demonstrate that the house detection is satisfactory enough.

The future work will be connected with the following improving and testing of this algorithm.

References:

1. Stepanov A, Zheltov S. and Visilter Yu., "Shape analysis using Pytiyev morphologic paradigm and its' use in machine vision". - SPIE vol. 2350, 1994.
2. Houle D. and Malowany A., "Object recognitions using fast Adaptive Hough transform", IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, June, 1-st-2nd, 1989.
3. Huertas A, Bejanin M. and Nevatia R., "Model Registration and Validation". - in "Automatic extraction of man-made objects from aerial and space images"/ ed.by A.Gruen at al, Basel: Birkhauser, 1995.
4. Muller W. and Olson J., "Automatic matching of 3D-Models to Imagery". - in "Automatic extraction of man-made objects from aerial and space images"/ ed.by A.Gruen at al, Basel: Birkhauser, 1995.