MULTI-SCALE IMAGE ANALYSIS and APPLICATIONS in FRAMEWORK of the INTAS 96-785 PROJECT

H.J.A.M. Heijmans
CWI, Amsterdam, The Netherlands
V.S. Kirichuk
Institute of Automation and Electrometry, Novosibirsk, Russia
V.A. Kovalev
Institute of Engineering Cybernetics, Minsk, Belarus
E.J. Pauwels
K.U.Leuven, Leuven, Belgium
M. Petrou
University of Surrey, UK
V. I. Tsurkov
Computing Center, Moscow, Russia
A.V. Tuzikov
Research-Engineering Center of Informational Technologies, Minsk, Belarus

The necessity of multi-scale methods for image analysis has been recognised since the emergence of the field of image processing. Processing images at different scales has many benefits as it gives additional information about the essence of the percept which we want to understand. Today, there exists various alternative approaches: linear and nonlinear image scale-spaces, morphological imaging techniques, quadtrees, pyramids, wavelets, fractal-based techniques, etc. The goal of the current project is to contribute to this important research area, and more ambitiously, to explore their interrelationships. Furthermore, we want to apply such methods to various concrete problems, in particular the analysis of medical images (texture analysis, segmentation).

An important application area where multi-scale methods can play an important role is Content-Based Image Retrieval (CBIR). Given the exponential growth of image and multi-media material in large databases and the Internet, there is increasing interest in search-engines that are able to retrieve images on the basis of image-oriented queries and visual clues. (Queries like: "Find images of deserted tropical beach" or "Find images that look like this one"). As compared to more traditionally oriented databases (where one looks for an exact match with the query-item), image databases are searched for objects similar to the query-item. Research has shown that more advanced techniques for automatic image analysis and more sophisticated ways of shape recognition will be necessary to perform such tasks.

Research activities

University of Surrey, UK
Institute of Engineering Cybernetics, Minsk, Belarus

1. Wavelets and 3D texture analysis.

Two approaches to the characterisation of 3D textures were studied. One based on gradient vectors and one on generalised co-occurrence matrices. They are investigated with the help of simulated data for their behaviour in the presence of noise and for various values of the parameters they depend on. They are also applied to several medical volume images characterised by the presence of micro-textures and their potential as diagnostic tools and tools for quantifying and monitoring the progress of various pathologies is discussed. The gradient based method appears to be more appropriate for the characterisation of micro-textures. It also shows more consistent behaviour as a descriptor of pathologies than the generalised co-occurrence matrix approach. Results are presented in [1].

2. Content-based image retrieval.

Multimedia documents are different from traditional text documents because they may contain encoding of raw sensorial data. This fact has severe consequences for the efficient indexing and retrieval of information from large unstructured collections (e.g. WWW) since it is very difficult to automatically identify generic meanings from visual or audible objects. Two sub-problems of content-based image retrieval were studied at this stage, colour image retrieval and shape retrieval. Generalised co-occurrence matrices were employed in both cases.

Colour image retrieval. Our method of colour image retrieval is based on so-called colour co-occurrence descriptors that utilise compact representations of essential information of the visual content [2]. The set of descriptor elements represents "elementary" colour segments, their borders, and their mutual spatial distribution on the image frame. Such representation is flexible enough to describe image scenes ranging from simple combinations of colour segments to high frequency colour texture equally well. At the retrieval stage the comparison between a given query descriptor And the database descriptors is performed by a similarity measure. Image descriptors are robust versus affine transforms and several other image distortions. Basic properties of the method are demonstrated experimentally on an image database containing about 10,000 colour images.

Shape retrieval. We have suggested a method that exploits the orientation tokens that characterise a shape or an image for object identification [3,4]. The method is appropriate for identifying objects from their silhouettes or grey level texture. The approach was tested with the help of a large database of marine animals (mostly fishes). Pictorial queries by shape were answered successfully using only the L2 norm for histogram comparison. The accuracy of the answer can be improved if more features are computed from the histogram. Some example features for this purpose were proposed. On the other hand, the method was successfully tested for the identification of textures in leaves, using as feature a measure of the "order" that appears to be present in a texture. It was proposed that this can be expressed by the ratio of the number of orientation tokens that are more or less parallel to each other over the number of tokens that form larger angles with respect to each other. The pairs of tokens considered are at fixed distance from each other. For a quick search of an image database this produces an efficient method of sorting images "at a glance". For a more detailed description, one could use many different ranges of relative distance and thus produce a 2D angle-distance co-occurrence matrix for the description of the object. The use of histograms for such a task is ideal as histograms are rotation and translation invariant. On the top, they are scale invariant as well, because we consider all possible areas of tokens and bin their distances in a fixed number of bins, independent of the actual range of distances contained in each bin. Finally the use of orientation tokens is independent of any illumination and grey level changes, something of major importance when scanning a collection of images that might have been captured under varied conditions. As a corollary of this investigation, we have shown some evidence that the MRI images of healthy subjects show more organised textures than the MRI images of patients suffering from various brain disorders. Results are published in [1,4].

3. Non-rigid body registration of 2D/3D images.

A novel approach for 2D/3D image registration based on non-linear elastic deformations and global optimisation is proposed. Suitable cost function containing similarity, deformity, and discrepancy terms is suggested and experimentally studied. The terms of cost function may be used as measurements of the degree of deformation, necessary to bring two images in registration, and also as quantifiers of the evolution of various changes like as pathology development in medical images. Results are published in [5,6].

Computing Center, Moscow, Russia

The analytical model of the edge in the image improvement problem by means of anisotropic diffusion is considered. In 1990 parabolic type differential equations in partial derivatives were proposed by P. Perona and J.Malik [7] for two-dimension image processing. The main idea of this method is the following. As is well known, in the case of a problem with initial conditions for diffusion (thermoconductivity) equation the tendency of noise smoothing appears due to some specific properties of parabolic equations. Thus noise can be reduced in image processing. However, according to the same property of parabolic equations the image edges with infinite initial gradients would be also smoothed. This would lead to the blurring of images. In order to remove this drawback, P.Perona and J. Malik proposed to use a diffusion coefficient, which depends upon the gradient. It is important that the diffusion coefficient should tend to zero as the derivatives become infinite. In other words, image noise should be eliminated while preserving infinite gradients, i.e. image edges. The Perona-Malik problem has put forward a number of publications. Among them regularisation is used in F. Cotte, P.-L. Lions, J.-M. Morel and T. Coll [8] and L. Alwarez, P.-L. Lions and J.-M. Morel [9]. The partial derivatives are represented as convolutions with Gaussians, which contain a small parameter. With this parameter tending to zero, the convolution tends to partial derivative. All mathematical transforms and numerical calculations are performed with the small parameter fixed (e. g. constant) and the solution is regular. In [9] it was stated that the limit solution is Lipschitz-continual for continual initial conditions as the small parameter tends to zero. Suggestion is made [9] that a divergent form of the Perona-Malik model may result in Adamar instability. That is why undivergent form of parabolic equations is proposed [9].

At this work the following results were obtained. Semi-similar solutions of anisotropic diffusion equations are found in one-dimension problem when a diffusion coefficient is a power function of the derivative. These solutions correspond to preserving of the infinite gradient with time. They are referred to as analytical edge model of anisotropic diffusion. It is found that the edge is preserved if the derivative power in the diffusion coefficient is more than 2. Otherwise, regular solutions take place. It is shown that the initial condition problem of the divergent Perona-Malik model is incorrect if the power degree number is more then two. We obtain that edge is a singular function of Gelder-continual class on the contrary to the Lipschitz-continual type limit solutions from [9]. An existence theorem for the weak (generalised) solution of the problem with initial conditions from Gelder-continual class is proved. Our prove is based on obtaining internal a-priori estimates for the first derivatives (or Lipschitz constants) of so-called potential functions, preserving regularity on edges. The weak solutions coincide with strong ones at points in which the diffusion coefficient is not equal to zero. At the edges the solution existence is considered as fulfilling of integral relationships with probe functions. A numerical experiment was conducted in which the initial edge considered as a Gelder-continues class function was perturbed by different types of noise. Results are the following. The edge is preserved and the image noise is reduced, thus Perona-Malik problem is solved. Semi-similar solutions of the anisotropic diffusion are obtained and analysed. They depend upon the ratio of a space co-ordinate to the square root of a time co-ordinate. Theorems on the asymptotic tending of the initial condition problem solutions to the solutions mentioned above are formulated in various topologies. Filter development on the base of the analytical edge model needs utilisation of discrete step numerical algorithm. Implicit discrete schemes of parabolic type equations are used in one-dimension calculations. Edge localising is performed with an interpolation procedure.

Research-Engineering Center of Informational Technologies, Minsk, Belarus

Research directions:

- Granulometries, pattern spectrum, and the opening transform;
- Morphological image analysis (shape similarity and symmetry measures, fast morphological algorithms, connected operators, texture classification);
- Content-based image retrieval.

Obtained results.

1. Similarity and symmetries of arbitrary polygonal shapes were investigated [13,14,15]. Notion of symmetry measure was introduced. Contour representation known as a turning function is used to evaluate symmetry measure. This representation is suitable for simply connected compact objects. Fast algorithms for similarity and symmetry measures can be implemented using such a representation. Analysing such functions we can compute a measure of reflection symmetry (with position of the best symmetry axis) or a measure of rotational measure (for different orders of rotations simultaneously). For affine symmetry the main idea consists in using canonical form normalisation in such a way that affine (skew or skew rotation) symmetry of original shape is equivalent to usual (reflection or rotation, respectively) symmetry of its canonical form. Algorithm of finding similarity and symmetry measure using turning function was implemented in C++ and tested for different objects and transformations (reflections, rotations, or affine transformations). The influence of noise to symmetry measures was tested also. User interface for this software was implemented in Java.

2. Symmetries of 3D convex shapes were investigates [16]. Symmetry measures based on Minkowski and Brunn-Minkowski inequalities for volumes and mixed volumes of convex objects were introduced for evaluations of symmetry degree. For the case of convex polyhedra functionals used in these measures have a nice property of piecewise concavity which allows to reduce the complexity of optimisation problems. For reflection symmetry of polyhedra the problem is reduced to checking of the functional value for the critical rotations of polyhedra, number of which seems to be finite. For rotation symmetry of order 2 the problem is the

same as for reflection symmetry. For rotation symmetry of orders more than 2 no optimisation method is proposed. The case of mirror rotation symmetry is equivalent to finding similarity measure of two polyhedra.

Implementation of polyhedra similarities and symmetries is now in progress. Several program components including finding of critical rotations for fixed axis are already written and tested using Matlab system.

Institute of Automation and Electrometry, Novosibirsk, Russia

Relief reconstructing and mobile objects detecting by sequences of images.

The data of continuous synchronous observation of analysed scene from several satellites are intended to be used with the purpose of detecting of dynamic processes occurring on the Earth surface and at the adjacent atmosphere layers (volcano activity, tornado, typhoons, explosions, conflagrations, etc.). The observing simultaneously and from different points provides selectivity of the atmospheric phenomena in the depth direction; the observing continuously enables to reveal their dynamic features.

Registration system geometry is not definite completely in the common case. So far, the problem of estimation of cameras orientation should be solved to restore the structure of the scene observed. Most approaches are based on relations between co-ordinates of corresponding points of projected images. Standard estimating fundamental matrix methods are not efficient because of cameras with small angles of view are used to get high spatial resolution. A great number of images in sequence observed and the knowledge about mathematical model that describes dynamics of the system geometry gives us new opportunities in sequences combined processing but requires new methods of data processing.

Two cases are investigated within the framework of the problem: a) the registration system is fixed in respect to a scene (geostationary satellites); b) the registration system is mobile (orbital systems).

The main purpose is to detect extremely mobile, slightly contrast objects when analysing data being received from geostationary system. Its complexity is conditioned by an image complexity -- an image contains stationary background (response from the Earth surface), quasistationary part (response from clouds), response from an objects and a noise. Moreover the intensity of response from an objects is much lower then both the intensity of response both from clouds and from the Earth surface. Besides there exist little shifts of registration system. Consequently every preventive factor should be reduced extremely to achieve good results in detecting an objects and their spatial location.

Thus, to solve the general problem one should solve several separate ones:
- to bind the sequence to some fixed system of co-ordinates using a correlative criterion;
- to carry out interframe processing to reduce powerful responses from the Earth surface and from clouds. The algorithm of interframe processing is based on monosemantic expanding of every current image with a set of binary images (basis functions), approximating an every consequent image with the functions and, finally, getting a corresponding differential image that does not contain stationary the background;
- to select (by means of linear and non-linear filtering) the points probably belonging to the object;
- to calculate an area of spatial location for every point detected for every sequence;
- finally, to combine results of processing for all sequences analysed and to select the points that areas of spatial location are intersected. The points are supposed to belong to the object(s).

When a registration system is mobile it is necessary to estimate parameters of the system (as a rule, cameras orientation is indefinite) under condition of extremely small angles of view (~1°), and to find corresponding points under conditions of great difference in angles of observation (up to 90°) and of great distance from the scene observed (~1000 km).

There are algorithms listed below designed to solve the problem:
- of finding corresponding points on the base of modified, geometrically tuneable to projective deformations, correlative and morphological criteria.
- of estimating geometry both of cameras location and of scene on the base of combined testing sequence images and solving essentially overdetermined system of algebraic equations.
- of restoring relief by sequences of images on the base of sets of corresponding points using LMS method.
- of global minimising the functional with the definite geometry of sequence and without finding of corresponding points.

Obtained results:

1. The algorithms of combined processing of sequences obtained from geostationary satellites are designed aimed to detecting extremely mobile, slightly contrast objects. It is efficient under conditions of time instability and powerful background signal (false alarm probability is $5*10^{-5}$, detection probability is 0.95).

2. The algorithm of restoring relief from a sequences of images under conditions of indefinite cameras observation angles is designed. Its accuracy is examined on models and tests. It achieves 65 meters when reconstructing relief of Erebus volcano by 24 images obtained from the distance of 1100 ÷ 1500 km (relief height ~0÷4 km) .

3. The algorithm of surface reconstruction by long (~100 images) sequence of images and difference of cameras observation angles 30° ÷ 40° is developed. Its accuracy achieves 1/n where «n» is number of images in the sequence. The algorithm uses a procedure of global minimising of the selected functional instead of extremely unstable procedure of corresponding points finding.

CWI, Amsterdam, The Netherlands

There exists increasing interest in the development of tools that consider images at different scales of resolution. In this respect pyramidal image structures are of particular interest. A well-known instance of such a structure is the wavelet transform, but there exist others, such as the Gaussian pyramid. In collaboration with Prof. J. Goutsias (Johns Hopkins University, Baltimore), Heijmans is investigating pyramid structures based on morphological operators [17,18]. One of the ambitious goals of this project is to make a systematic study of nonlinear (morphological) wavelets.

Connected operators form a special class of morphological operators which act on the level of flat zones of images rather than on individual pixels. As a consequence, such operators can delete edges, but cannot change them, neither their shape nor their location. As a result, connected operators are well-suited for many imaging tasks, such as segmentation, filtering, and coding. The aim of this project is to develop a consistent mathematical theory for connected operators, and to investigate their implementation using a tailor-made graph structure. In 1998, the activities in this project have been very limited, but they will be taken up again in 1999.

Scale-spaces are a modern bottom-up tool in computer vision. In scale-space theory one embeds an image into a continuous family of gradually smoother versions of it. Increasing the scale should then simplify the image without creating spurious detail. The prototype example is the Gaussian scale-space, the construction of which is based on filtering the image with Gaussians of increasing width. Alternatively, one can view the resulting family of images as a solution of a linear diffusion equation. Recently, various nonlinear scale-spaces have been constructed, including those based on morphological operators. Together with van den Boomgaard of the University of Amsterdam, Heijmans has proposed a new axiomatic framework for scale-spaces based on algebraic concepts. This framework includes various existing examples, including Gaussian and morphological scale-spaces. A technical report is in preparation.

Katholieke Universiteit Leuven, Belgium

Within the framework of this project Content-Based Image Retrieval (CBIR) has been identified as an application that serves as a testbed for the theoretical elaborations of the multi-scale paradigm in image analysis. We recall that CBIR deals with the problem of automatically retrieving images that are perceptually similar to a given query-image from a large digital databank.

As a starting point we therefore looked at number of existing systems, such as QBIC (IBM), Virage, Photobook (MIT), etc,... to clearly identify the current state-of-the-art. We found that most of these systems are guided by pixel-based similarity measures. However, extensive experimentation over the last few years has shown that matching natural images solely on that basis of global similarities is often too crude an approach to produce satisfactory results. What is required is some form of perceptually relevant segmentation that allows one to identify a (small) number of salient and semantically meaningful image-regions which can then serve as the basis for more discerning region-based matching, e.g. it stands to reason that when it comes to perceptual matching, the foreground in an image is far more important than the background.

We therefore concentrated in this first part of the project on a generic segmentation-methodology that, using a number of basic features such as colour and texture, divides an image in a small number of perceptually salient regions, which can then be used as the starting point for similarity matching. In essence the methodology is very simple: after mapping the pixels of an image into a feature space (e.g. colour-space), we can proceed to cluster the data in this feature space. Identifying the thus-obtained clusters in the original image, yields a segmentation.

Hence, from an abstract point of view, segmentation and perceptual organisation can be interpreted as a problem of selecting appropriate features, followed by cluster-detection in feature-space. In fact, we can tighten up this argument even further since both steps are but two aspects of the same problem, as a particular feature-space is deemed appropriate whenever it shows pronounced clusters. Indeed, if mapping the pixels into the feature-space lumps them all together, this particular set of features is obviously of little use in defining perceptual saliency.

However, clustering problems commonly encountered in CBIR-applications are particularly challenging as mapping the images to feature-spaces often produces very irregular and unbalanced data-clouds. Furthermore, given the fact that segmentation and region-extraction should proceed automatically, we cannot assume that prior knowledge about the number of clusters or their shape is available. We have therefore developed a robust and versatile nonparametric clustering algorithm that is able to handle the unbalanced and highly irregular clusters encountered in such CBIR-applications. The strength of our approach lies not so much in the clustering itself, which is based on non-parametric density-estimation, but rather in the definition and use of two cluster-validity indices that are independent of the cluster-topology. By combining them, an optimal clustering can be identified, and experiments confirm that the associated clusters do indeed correspond to perceptually salient image-regions [19].

For more details and examples, we refer to our website http://www.esat.kuleuven.ac.be/~frederix/segmen-tation.html .

In the second part of our work we focussed on the problem of classification. The relevance for CBIR is clear: once prototype-images have been put into different groups, new images need to be classified into the existing classes so that they can be indexed and identified. There is a huge literature on classification and Support Vector Machines (SVM), one of the more recent methodologies, has enjoyed a lot of attention in recent years. It uses the training data to construct an optimally separating hypersurface between two classes. This surface is represented by a small number of "support vectors" that are used to classify new data. If the data are linearly separable, these support vectors can be determined directly. If this is not the case, SVM sidesteps the non-linearity by mapping the data into a higher dimensional space where a linear separator can be found. However, the nature of this higher dimensional space is highly data-dependent and finding an appropriate transformation is often difficult and left to the user. We have developed a new method that finds support-like vectors (for which we coined the name "sentinels") in the original data-space (so there is no need to find an appropriate transformation to a higher dimensional space). It is similar to LVQ and k-means in that it is based on attraction between data-points. However, unlike these aforementioned methods, it identifies datapoints near the boundaries of the classes, which can then be used for classification. A paper describing this algorithm and its performance is in preparation.

References

1. V.A. Kovalev, M. Petrou and Y.S. Bondar. Using orientation tokens for object recognition, Pattern Recognition Letters, Vol. 19, No. 12, pp. 125-132, 1998.
2. V. Kovalev and S. Volmer. Color Co-occurrence descriptors for querying-by-example, The International Conf. on Multimedia Modeling (MMM'98), Oct. 12-15, 1998, Lausanne, Switzerland, IEEE Comp. Society Press, pp. 32-38, 1998.
3. V.A. Kovalev, M. Petrou and Y.S. Bondar. Texture anisotropy in 3D images, IEEE Trans. on Image Processing, Vol. 8, No. 3, pp. 346-360, 1999.
4. V.A. Kovalev and M. Petrou. Image retrieval by object shape using co-occurrence matrices, 5th Int. Conf. On Pattern Recognition and Information Processing (PRIP'99), Minsk, Belarus, May, 1999 (in Russian, in press).
5. V.A. Kovalev, A method of elastic exponential deformations for image registration,(in Russian, in press).
6. V.A. Kovalev, Registration of 2D and 3D images of non-rigid objects, (in Russian, in press).
7. P. Perona and J. Malik Scale-space and edge-detection using anisotropic diffusion. IEEE Trans. on Pattern. Anal. and Mach. Intel. Vol. 12, N 7, pp. 629-639, 1990.
8. F. Cotte, P.-L Lions, J.-M. Morel and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion I.SIAM J. Numer. Anal., Vol. 29, N1, pp. 182-193, 1992.
9. L. Alwarez, P.-L Lions and J.-M. Morel. Image selective smoothing and edge detection by nonlinear diffusion II. SIAM J. Numer. Anal., Vol. 29, N 3, pp. 845-866, 1992.

10. V.S. Kirichuk, V.P. Kosykh, G.I. Peretjagin. Extrinsic and Intrinsic Parameters Estimation for Stereo-Vision System with a Common Observation Point (Long-Focal Approximation). 4th All-Russian with invited foreign participants Conf. on Pattern Recognition and Image Analysis (ROAI-98), Akademgorodok, Novosibirsk, Russia, Oct 11-18, 1998 (in Russian).
11. V.P. Kosykh, G.I. Peretjagin. Recovering the 3-D Structure of Surfaces on Sequence of the Images Obtained from a Moving Camera Directed to the Same Point of Scene. 4th All-Russian with invited foreign participants Conf. on Pattern Recognition and Image Analysis (ROAI-98), Akademgorodok, Novosibirsk, Russia, Oct 11-18, 1998 (in Russian).
12. V.S. Kirichuk, V.P. Kosykh, G.I. Peretjagin. Scene structure and camera motion estimation using long image sequences (longfocal approach). Pattern Recognition and Image Analysis (accepted).
13. S. Sheynin, A. Tuzikov, D. Volgin. Computation of Symmetry Measures of Polygonal Shapes. Proceedings of CAIP'99, Ljubljana, Lecture Notes in Computer Science, Springer, 1999.
14. S. Sheynin, A.Tuzikov, and D. Volgin. Symmetry measures of polygons based on the turning function. (Russian) – Pattern Recognition and Information Processing, Minsk, 1999, Minsk-Szczecin, 1999, Vol.2, pp.90-94.
15. S. Sheynin, A.Tuzikov, and D. Volgin. Symmetry measures computation for polygonal shapes. Pattern Recognition and Image Analysis, N.3, Vol.9, 1999.
16. A. Tuzikov and S. Sheynin. Symmetry measures for 3D convex shapes. In: D. Chetverikov, T. Sziranyi (eds). Fundamental Structural Properties in Image and Pattern Analysis 1999, International workshop, Budapest, Österreichische Computer Gesellschaft, 1999, band 130, pp.77-86.
17. J. Goutsias and H. J. A. M. Heijmans. Nonlinear multiresolution signal decomposition schemes. Part 1: Morphological Pyramids. IEEE Transactions on Image Processing (to appear).
18. H. J. A. M. Heijmans and J. Goutsias: Multiresolution signal decomposition schemes. Part 2: Morphological Wavelets. CWI Research Report PNA-R9905, Amsterdam 1999.
19. E.J. Pauwels and G. Frederix: Finding salient regions in images: Non-parametric clustering for image segmentation. Computer Vision and Image Understanding. Special Issue on Content-Based Access of Image and Video Libraries, June 1999, 16 pages.