FRAMES BASED ANALYSIS OF MULTISENSOR IMAGE SEQUENCES

Yury V.Morzeev, Yury V.Visilter, Sergei Y. Zheltov, Alexander A. Stepanov
(State Research Institute of Aviation Systems, Vicktorenko 7, Moscow, Russia)
e-mail: yury@fenix.niias.msk.su

KEYWORDS: multisensory data processing, object-oriented programming, frame net, data fusion, image sequences.

ABSTRACT

At present day we can see the increasing of interest to the problem of combined analysis of the multisensor data. One of the prospect way to solve the fusion problem for this heterogeneous information for the improvement of the informational-perceiving characteristics of the designed system is the developing of the high-accuracy multisensor data fusion algorithms. This paper dedicated to generalize the results, which we achieved in the field of multisensor data fusion analysis.

1. INTRODUCTION

The appearance of the new 2D-sensors generation (e.g., video, IR, LL, etc. with a lot of modifications) increase the interest to the problem of combined analysis of the multisensor data [1]. The developing of this new real-time measuring systems and registration of the 2D-images methods makes it possible to process the time-invariant images, which have different physical nature.

The enhancement of informational perceiving characteristics of the designed sensor systems can be achieved by solving of multisensor data fusion problem. Hereupon, the final system could be designed to solve one of two principle problems:

- the complex support to help the operator in decision making problem in the case of the half-automatic targets selection systems;
- the automatic decision making in the automatic targets selection systems.

Due to these problems, we would like to generalize the results, which we achieved in the field of multisensor data fusion analysis. In this paper the new object-oriented approach to multisensor data processing will be discussed. This approach presumes any processing scheme to be represented as a *network of software frames* [2]. The interaction between objects in such network is provided by means of the message transmission between frames in accordance with some logic rules. The designed technology provides the automatic analysis of multisensor image sequences.

2. BASIC PRINCIPLES FOR MULTISENSOR IMAGE SEQUENCES PROCESSING

In this section we want to discuss the generic *Multi-Sensor Data Processing Framework* (MSDPF) that, being implemented using our software frames, can support the most wide spectrum of possible processing schemes [3]. The most strong limitation of our framework is a *modularity* assumption. This term, in our comprehension, means that the mapping process of input data to output data is structured into the set of sequential or parallel procedures with input/output dependence. Some buffer data structures that are the output of some procedures and simultaneously the input of some other procedures express this dependence. It is easy to see that only the structured processing schemes can have the adequate frame-based representation. It is obvious that the set of control frames is not problem-dependent. Thus, we have to specify the set of data frames and the set of processing frames. Due to data-driven character of our approach, the set of data frames must be firstly defined and the set of processing frames will be defined later, on the basis of known data frame types. For simplicity, we shall describe our MSDPF in terms of data types and procedure types.

2.1. The levels of data abstraction

It is well known nowadays, that there is the customary conception of *data processing levels* in the field of data fusion problem. It usually corresponds to the sequential modular scheme (see fig.1) [1].

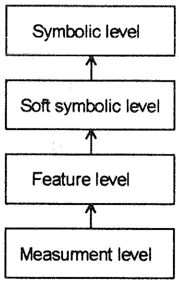

Fig. 1. Multi-sensor data processing semantic levels.

The arrows in this scheme show the direction of information movement: from the *measured signal* through the *feature-based descriptions* to the resultant *symbolic description.* Last years, the symbolic level is often considered as a union of two sub levels: the *soft* (probabilistic or fuzzy) *description* level and the proper *symbolic* level corresponding to the final decision making about the elements of the observed scene. One supposes that the processing starts from the separate processing of measured data from each sensor and the fusion takes place at one of the highest levels of this scheme. Since the fusion has being done, the further processing of the fused data is executed if it is required.

Thus, we have the *interlevel* (down to up) and *on-level* transformations of data. The interlevel transformations extract the usable information from data of lower level and the on-level procedures realize the proper data fusion. Today the most attractive fusion levels are the symbolic and the feature levels. The measurement (image, pixel) level is usually rejected because it is not easy to provide the accurate co-registration of multi-sensor image data. However, it takes place in the *remote* sensing case where the co-registration is accurate enough. So, the measurement level must be also of consideration as a fusion level.

We agree all mentioned points but the introduced terms are too generic and uncertain to specify the proper data types and processing procedures. To define the required set of frame types we need to analyze the problem more detail. The first new point we propose to consider is that any data at any processing level represent a *structure* of some *elements*. According to this point of view, one can say that fig.1 describes the levels of data abstraction for elements but not for structures. The analogous scheme for structures may have, for example, the following form (fig.2):

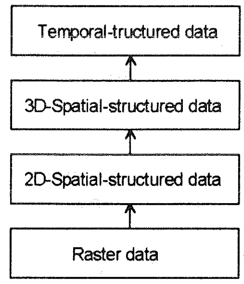

Fig. 2. Data organization levels.

The basic level of this scheme corresponds to the sensor-generated level of data organization. Second level corresponds to the segmented image data. The third level includes the 3D-scene descriptions and the top level of this scheme deals with the time-varying 3D-world. In general, this scheme is not right because it mixes two different types of data organization: *spatial* and *temporal.* These organization types are independent from each other. So, we have to consider three spatial organization levels and three different levels for time-varying data: raster, 2D-structured and 3D-structured. However, it is the *long-range remote sensing* case where the 2D-temporal data (raster or structured) are seldom of use. So, we adopt the scheme above (fig. 2).

All of levels (fig. 2) are the abstraction levels too, however, there is no any correlation between these levels and levels mentioned before. For example, the objects of 3D time-varying world description may be characterized by the feature vectors and, conversely, symbolic data may be stored in the raster form. Thus, we have the 2D-space of combinations of possible data structures with different element types. So, any data type can be described with the use of two "co-ordinates" - *semantic level* and *organization level*. Let's note that the fusion may also be executed at any level of scheme 1 and simultaneously at any level of scheme 2. So, now we can not more characterize the fusion operators as on-level operators in scheme 1. The fusion procedure must preserve the both of data type co-ordinates.

The introduced 2D-space is just enough to specify the set of data frames. Nevertheless, one more additional point must be outlined before. It is the rare case when the object at feature semantic level is characterized by only one feature. Usually the feature level presumes the feature vector of some dimension to be corresponded with each element of data structure. Analogously, the soft symbolic description associates any structural element with the vector of fuzzy measures (probabilities) of hypothesis. The dimension of this vector is equal to the number of object classes known for the proper system. Finally, the measurement element can be also the vector of some dimension, e.g. the TV-signal provides the three-component pixel value RGB. Thus, the *dimension of element* must be considered as a third co-ordinate of data types. Analogously, the *dimension of structure* must be the fourth data type co-ordinate. However, nobody thinks that the images of different size are the different data types. So, let's adopt the following agreement: the *dimension* will be the binary characteristic taking its value on the set of **{one, any}**. It means that any element is presumed to be either the one number (symbol) or some list (vector) of arbitrary length. This agreement makes the software implementation of our framework much more easy [1].

2.2. Procedure types

All the necessary procedure types have been discussed in [1]. That's why we'll only mark the principle moments. As mentioned above, the processing frames have some input/output information links. So, some *input set* and *output set* must be determined for each principal processing frame type. The input set description includes the number of input data with their types. The output set description includes the number of output data with their types. In this section we shall consider the structure dimension as a fourth co-ordinate of data type.

Generally, it is possible to define the mapping of any input set onto any output set. However, it leads to the set of procedures of unobservable size. So, we need to introduce some additional constraints.

The first important constraint concerns the possible content of input and output sets:

any input(output) set can contain only data of the same semantic level and same organization level.

The second important constraint we refer as an ***Exclusive Modularity Principle:***

any mapping of any input data set onto any output data set can be represented through a combination of on-level and exclusive inter-level procedures only.

The third constraint declares that:

any inter-level procedure satisfies the condition of $N_{in}=1$ *and* $N_{out}=1$ (where N_{in}, N_{out} - number of inputs and outputs, correspondently).

In particular, it means that the fusion procedures change neither organization nor semantic level of data representation.

The fourth important constraint we adopt in the form of the ***Cumulative Fusion Principle:***

any structure fusion procedure can be represented as a combination of number of the pair-wise fusion procedures.

It is very powerful assumption that is not true for some usable fusion approaches. However, the cumulative fusion proposal makes it possible to design the required fusion scheme using the unified fusion frames (with two input and one output links) for any possible set of sensors.

The last constraint of our framework closes the set of on-level procedures:

only two types of on-level procedures are available: filters and pair-wise fusion procedures.

Now, after this preliminary discussion, we are ready to outline the developed set of procedure types.

2.2.1. Exclusive inter-level procedures.

There are two types of the exclusive inter-level procedures: procedures that transform data with level increasing and procedures that transforms data with level decreasing. When the level of data abstraction increases, the extraction of information occurs. This case takes place immediately during data processing and fusion. The converse case corresponds to the data modeling process. We shall not consider any modeling here.

A. Semantic inter-level procedures. Procedures of this class preserve the structure while updating the elements. The most important procedures are:

- *feature extraction* - calculates some feature for each element using measurements; there must be a set of such procedures to obtain the required feature vectors;
- *feature-based classification* - assigns the symbolic (class) labels for each element using its features; may be of soft or hard type; may the Bayesian classification, cluster analysis and so on;
- *object detection* - assigns the symbolic (class) labels for each element using the matching techniques at the measurement level; may be of soft or hard type; may use the Bayesian classification, correlation, model-based methods and so on; there must be a set of such procedures if a set of known objects is given;
- *decision making (recognition)* - assigns the unique symbolic (class) label for each element using the soft (probabilistic, fuzzy) or hard evidences from different sources.

B. Structural inter-level procedures. Procedures of this class preserve the element types while updating the structural organization. The most important procedures are:

- *segmentation* - transforms the raster data into regions and stores them in some structure; may use the contour-based, region-based, texture-based, relaxation and other techniques;

* *scene restoration* - forms a scene description, based on segmented structure; may be model-based or use some data base; extracts the spatial objects;
* *object tracking* - in the most broad sense, presumes the estimation of the element movement; requires to store the results of previous observations; may use the sequential image differences, correlation tracking, Kalman filtering, Bayesian estimation and so on; extracts the spatio-temporal events.

2.2.2. Pairwise fusion procedures.

Procedures of this class fuse two input structures of some type into the one output structure of the same type. There are two types of fusion: element fusion and structure fusion.

A. Element fusion procedures. The most important procedures are:

* *element theoretic-set union* - fuses two structures with corresponding element dimension of D1 and D2 into the resultant structure with element dimension (D1+D2); preserves the structure dimension; typically of use in fusion at the measurement and feature levels; results in creation of common space of measurements or features;
* *soft data fusion* - for each soft element fuses the soft characteristic vectors; preserves the both structure dimension and element dimension; may use Bayesian or Dempster-Schaffer reasoning;
* *evidence fusion* - fuses evidences to obtain the hard symbolic data; preserves the both structure dimension and element dimension; may use any evidence reasoning approach;

B. Structure fusion procedures. The most important procedures are:

* *structure union* - theoretic-set union of two structures; preserves the element dimension; of use in fusion of any structured data;
* *structure intersection* - theoretic-set intersection of two structures; preserves the element dimension; of use in fusion of any structured data;
* *rang fusion* - theoretic-set union of two structures with accumulation of number of evidences for each element and exclusion of elements with this number less than some rang (threshold); preserves the element dimension; of use in fusion of any structured data;

These procedures must be implemented for each of data structures supported by the proper system of multi-sensor data processing using the structure union, intersection or rang fusion.

2.2.3. Filtering procedures.

These are the on-level procedures with one input and one output. These procedures always preserve the element dimension of data. So, we need to consider the structure dimension as a characteristic feature of these procedures.

A. Dimension preserving procedures. The most important procedures of this type are:

* *element transforms* - transform the each element in the uniform way; typically of use at the measurement and feature levels, e.g. histogram transforms of intensity images or the mappings of feature spaces;
* *linear and non-linear transforms* - processes raster data in the spatial or frequent domain; the large set of such procedures can be implemented.

B. Dimension non-preserving procedures. The most important procedures of this type are:

* *geometric transforms* - transform any spatial data using some geometric model; the large set of such procedures can be implemented; may be of use to provide the co-registration conditions;
* *structure evolving* - processes the structure to obtain to better features or behavior, for example, 2D-segmentation by repetitive splitting and merging.

Thus, we have formed the complete theoretical framework for multi-sensor data processing. The necessity of implementation of some proper kind of filtering procedure is determined by the properties and required features of any certain multi-sensor application to be developed on the basis of this framework.

This theoretical basis allows to form generic frame-based software architecture for multi-sensor data fusion and processing of multisensor image sequences.

3. THE CAPTURING OF MULTISENSOR IMAGE SEQUENCES

In order to form the satisfactory and representative data base of image sequences the special experimental complex have been designed (fig.3). It includes:

* coherent LL-sensor;
* IR-sensor AGA-780 with the control system and monitor;
* high resolution video camera FTM-800;
* DISCON - the transmission device to transform the IR-signal to video standard;
* PC to store the IR and video images;

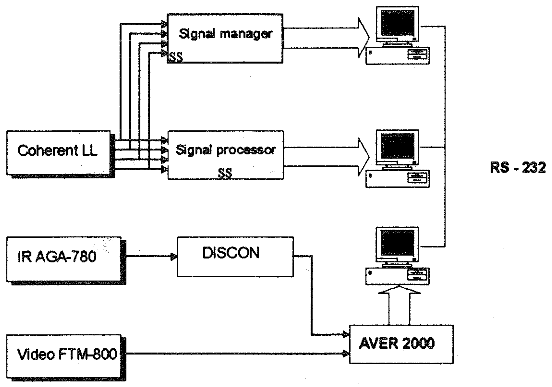

Fig.3. The functional scheme of experimental complex

We choose the experimental position on the height of 30m at the distance to the objects of 150÷2000m. As the example we choose the track ZIL-131. Some experimental results (TV, LL and IR images respectively) are shown on fig.4-5.

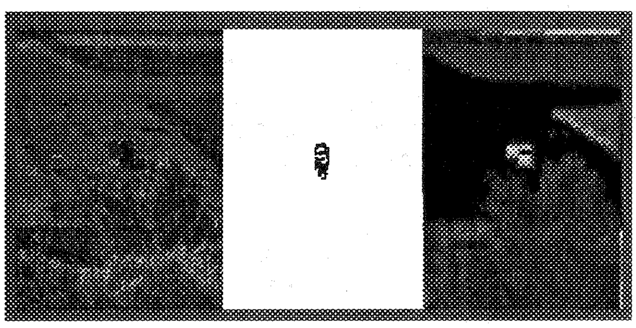

Fig.4 Truck at the distance about 500m

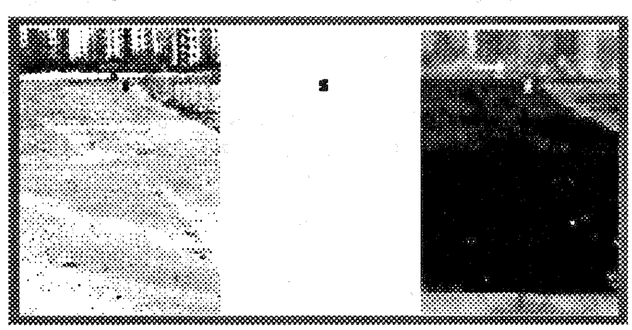

Fig.5 Truck at the distance about 1500m

4. MAIN RESULTS FOR FRAME BASED ANALYSIS OF MULTISENSOR IMAGE SEQUENCES

4.1. Base algorithms for multisensor image processing

In order to optimise the multisensor image sequences analysis on the level of processing we suggest several approaches to use the multisensor information to increase the reliability of target recognition in automatic manner:

- fusion on pixels level;
- fusion on the level of feature vectors;
- fusion on the level of soft-vectors.

For fusion on pixels level we have used Pytiev correlation theory and desiered the special algorythms [8]. Also, our software includes about 25 different features' extractor algorythms and designed the special approach to select the invariant feautures to the concrete problem. Due to the soft-vector level, there are two approaches to evidence fusion from different sensors. First of them is based on classical Bayessian theory, and the second one - modified Dempster-Schaffer approach, which was successfully used in our software[4÷7].

4.2. Software realization

The first subject of discussion on the problem of software realization is the determination of principle hardware platform. Any designer must find the compromise between requirement characteristics of the computer system and economical indexes.

The designed software is based on the conception of Frame Object-Oriented Programming [2]. This technique is based on the new type of programming objects - *"software frames"* (the term frame had taken from Artificial Intelligence). From the program point of view they are the generalizing of the well-known term "class" from the Object-Oriented Programming (OOP).

Frames are encapsulate links - special data structures, which determines the behavior of the frame. There are links of two types: "synchronies" and "information". Any frame action could be carry out as the reaction on the message from the other frame, which is connected with it by corresponding "input" synchronic link. Moreover, any external data are available for reading or writing only by "input" or "output" information links and only when the "frames-masters" are "agree" with it. Frames couldn't process independently or "see" any external data. Frames work only together.

In terms of OOP we could say, that any frame is the store place for the object, which contains the other objects, which are data, procedures, links, etc. So, frame, even if it hasn't definite members, has the possibility of access to them. Any copy of existing objects have to be placed in special "box" - slot - of frame structure.

Our approach suppose to present any processing scheme as a Frame Processing Set (FPS). FPS is an asynchronic set of soft-frames, which are connected by links (fig.6). Any FPS must include *Data Frames, Processing Frames and Control Frames.*

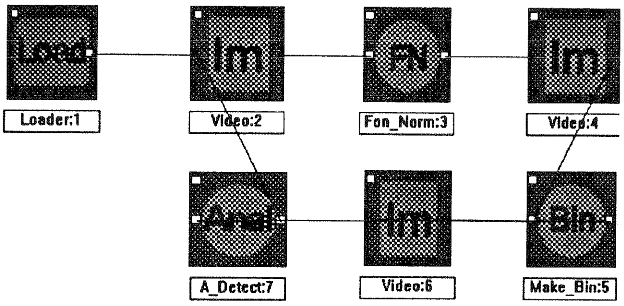

Fig.6. Example of typical FPS

In order to solve the problem of multisensor image sequences analyis there were designed about 30 different types of frames. Several of typical schemes with frames are shown on fig.7-8.

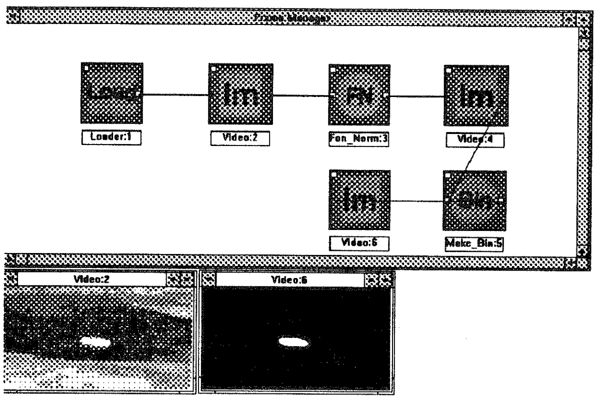

Fig.7. Background normalization (the kind of target detection), bynarization

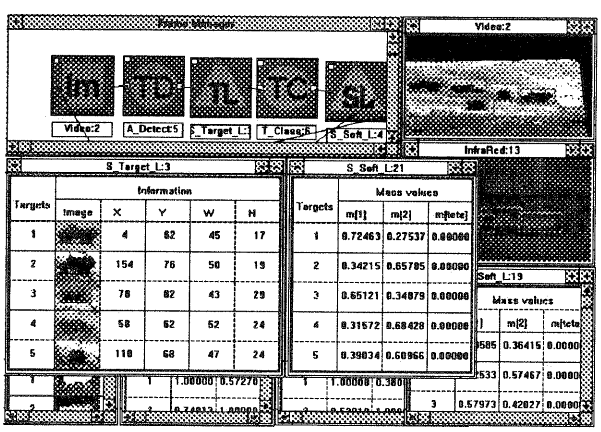

Fig.8. Fusion on the soft-vectors level

RESULTS AND DISCUSSION

This paper generalizes the achieved results of work in the field of multisensor data fusion analysis. We've discussed the generic *Multi-Sensor Data Processing Framework* (MSDPF) that, being implemented using our software frames. It can support the most wide spectrum of possible processing schemes. Frames encapsulate links - special data structures, which determines the behavior of the frame. In order to solve the problem of multisensor image sequences analysis there were designed about 30 different types of frames. The research in the field of fusion shows the efficient of designed algorythms because of increasing reliablity of targets recognition. Several approaches of multisensor information usage in order to increase the reliability of target recognition in automatic manner have been suggested.

REFERENCES

1. **Yu. Visilter, Yu. Morzeev, A. Stepanov, S. Zheltov** "Frame paradigm and object-oriented image processing for Photogrammetry", "St. Petersburg - Great Lakes Conference on Digital Photogrammetry and Remote Sensing '95", SPIE Proceedings, vol. 2646., 1995.
2. **Stepanov, Yu. Visilter, Yu. Morzeev, S. Zheltov** "The object-oriented frame approach to image processing and data management in the multisensory remote sensing", SPIE proceedings, vol. 2587, 1995.
3. **Yu.V. Vizilter, Yu.V.Morzeev, A.A. Stepanov, S. Yu. Zheltov** "Object-Oriented Frame approach for improved interfaces design for image processing systems", SPIE Proceedings, vol. 2588., 1995.
4. **Dempster** "Upper and Lower Probabilities Induced by a Multivalued Mapping", Annals of Mathematical Statistics, Vol. 38, pp. 325-339, 1967.
5. **Shafer**, "A Mathematical Theory of Evidence", Princeton Univercity Press, 1976.
6. **Ray Smith**, "A Bayessian approach to multisensor data fusion". Proc. of the SPIE. 1992, v.1699.
7. **Kim**, "Bayesian Inference Network: Applications to target tracking". Proc. of the SPIE. 1992, v.1698.
8. **Pytiev Yu.P.** "Morphological image analysis", 1983, pp. 1061-1064.