Some New Concepts and Key Techniques in Multi-Sensor Image Fusion

ZOU Mou Yan [1,2] and LIU Yan [2]
[1] *Institute of Electronics, Chinese Academy of Sciences*，
[2] *Graduate School, Chinese Academy of Sciences,*
Beijing 100190, P. R. China
Email：zoumouyan@mail.ie.ac.cn

An image can be understood as a Distribution of Attributes of Objects (DAO) in the image. Image fusion is a part of information fusion or association of attributes of objects on the basis of acquired images. From this understanding, image fusion is not a "pure" image processing problem. Rather, researchers should pay more attention to the underlying physical implication of images. In this paper we bring forward a few new concepts that would be essential in the theory and technique of multi-sensor image fusion. On this basis, a few key techniques are discussed relating to image fusion. The range of the related research work on multi-sensor image fusion remains large, multifold and challenging. More research is needed on the theoretical and conceptual models of image and data fusion.

Key words: Multi-Sensor Image Fusion, Image Registration, Image Segmentation, Feature Extraction, Image Interpretation, DAO, Association Database

1. Introduction

Image fusion is an important part of information fusion, which is the technology that combines several images of the same area or the same object under different imaging conditions (such as at different time, different weather conditions, different sensor types, different working waveband, different polarization modes, etc.) to implement the image processing task effectively，such as pattern or object identification, region segmentation and statistics, regional change description, etc. [Moser and Serpico 2009, Pastina and Spina 2009, Hizem et al. 2009, Wang et al. 2009, Qian et al. 2009, Bovolo et al. 2010]. These image processing tasks aim to serve for superior-level application in the statistical analysis and investigation of land resources and urban development [Unsalan 2009, Hedman et al. 2010], the statistic and investigation of agriculture and forestry and disasters [Petillot et al. 2010, Brunner 2010, Ma and Yang 2009, Licciardi et al. 2009], some military affairs, and more [Wu et al. 2009, Cartmill et al. 2009, Thomas 2008, Simone 2002, Gamba 2005, Karantzalos and Paragios 2010, Calhoun and Adali 2009, Gundimada and Asari 2009]. For different application purposes, image fusion involves common techniques as well as some different techniques and requirements [Yang and Li 2010, Looney and Mandic 2009, Copeland et al. 2010, Joshi and Jalobeanu 2010, Jun et al. 2009, Kumar and Dass

2009, Molch 2010, Monwar and Gavrilova 2009, Masini et al. 2009, Chen et al. 2010, Wan et al. 2009, Zhang et al. 2009, Li and Leung 2009, Xie et al. 2010, Poh et al. 2010, Celik and Ma 2010, Udelhoven et al. 2009, Jeffrey et al. 2009, Kalka et al. 2010].

One typical application of image fusion is as follows. Through a reconnaissance flight, someone found a tent in the forest from an aerial photo; at the same time by making use of an infrared imaging device he discovered a heat sources inside the tent; and furthermore a radar imaging apparatus reported that there existed metal objects inside the tent. Naturally, much complete information about the tent can be acquired if we combine these three pictures together than could be obtained from a single sensor alone.

In the field of land surveying and mapping, geologic surveying, agriculture evaluation, and weather forecasting, it is usually needed to synthesize one or more remote sensing images with an artificial geographical map to achieve different purposes. For example, a more clear and visual image interpretation helps the non-specialist to visually understand information expressed in the image, such as weather changing, disaster circumstances, regional distribution of various data, etc. However, from the view angle of image processing, this technique would be called as *image composition or image synthesis,* which is different from *image fusion* technically and conceptually [Kumar 1995, Flusser 2007, Goshtasby 2007]. Image synthesis is the technique that synthesizes various required contents of several images *with known content* to form a new image according to the application purpose. But the problem resolved by image fusion is just to identify the image content (identify objects, identify characteristics of the image region, identify image changes, etc.). Since a single image often cannot provide sufficient information for identification, we need to combine images acquired by multi-sensors or under different conditions to improve the accuracy of interpretation and the comprehensiveness of knowledge. In this way, image fusion is a research branch of data fusion [Pohl and Van Genderen 1998, Hall 1997], while image synthesis needs not to be built on these concepts.

The concept of information fusion appears earlier in a research report of the American military aspect in 1972. Suppose that there are several information sources will all take effect when an electrical system or weapon equipment acts. In such a case, we must find a method to synthetically determine an optimal action according to all information sources. Such method is called information fusion. This concept quickly spread to the command, control and decision systems and many other applied fields.

Generally, image fusion is performed at three different processing levels: Pixel, Feature, and Decision level [Pohl and Van Genderen 1998，Hall and Llinas 1997, Dong et.al. 2009, Simone et al. 2002, Vijayaraj et al. 2006, Smith and Heather 2005, Blum and Liu 2006]. ***For many years, researchers in our area popularly accepted such a definition: “Image fusion is the combination of two or more different images to form a new image by using a certain algorithm” [Genderen and Pohl 1994]. The number of published research reports on this topic is increasing year***

by year. However, the criticisms from the remote sensing application aspects should not be overlooked: in recent twenty to thirty years, no essential progress has been achieved yet by image fusion techniques for object identification, classification, and change detection. This fact impacts us to recheck the related research work. We feel that many researchers have paid their attention to the research of pixel-level image fusion, but less effort to the other aspects. Pixel-level image fusion is useful in many cases [Pohl and Van Genderen 1998, Zhang et al. 2009, Kumar and Dass 2009]. For multi-sensor image fusion, however, pixel-level fusion does not always meaningful. In a typical remote sensing application, for example, the results of image fusion may be exhibited by using artificial tags on a geographic map or several geographic maps with a few pseudo-colored areas. In fact, the artificial tags and the pseudo-colored areas are much more convenient and intuitive to describe the abundant attributes of objects. In such a case, the task of image fusion is to produce a set of numerical describers for objects, rather than a fused image.

The simple example mentioned previously has already concealed the basic meanings and requirements of image fusion technique. (1) A preprocessing step is required that implements the geometric alignment of several images acquired by different sensors. Such techniques are called as multi-sensor image registration. (2) The secondary preprocessing step is also needed that segments the commonly interesting region of images. (3) The third preprocessing step should solve the problem of extracting and describing the attributes or features of the concerned objects (targets or target regions) in every interesting region of images. (4) Finally, image fusion should be carried out that fuses information of attributes of interesting regions or/and the concerned objects and producing an image interpretation according to the applied requirements. It can be seen that every request of the above four aspects is challenging. The three preprocessing steps do not belong to the fusion processing itself, but they are vitally important for producing the desired fusion results.

The aim of this paper is to bring forward a few new concepts to the researchers in our area. These concepts are not fully consistent with the current view points on image fusion techniques. (1) “The combination of two or more different images to form a new image by using a certain algorithm” could not be the essential of image fusion. (2) An image can be understood as a Distribution of Attributes of Objects (DAO) in the image (See Section 2 below). The DAO depends on the physical properties of the objects as well as on the imaging apparatus and its work mode (wavelength or frequency, polarization, radiation or reflection, etc.), and the circumstance conditions. (3) The feature-level image fusion is just data association of objects. That is: the all attributes of an object extracted from all acquired multi-sensor images are associated to each other. (4) On the decision level, the associated data are used to produce the set of numerical describers of every object, according to the request of object identification, classification, and change detection, etc. To this end, certain decision rules are necessary.

The display of the fusion results and the related techniques are important and the display may be in various forms: an image with enhanced objects; an image with object tags; a geographic map with object tags; several registered images or several pseudo-colored geographic maps that can be displayed singly or compositely, etc. The conventional image processing techniques, such as geometric correction, scaling and resample, intensity and chromaticity correction, edge enhancement and filtering, etc., are important in image fusion [Pohl and Van Genderen 1998], but the related discussion will be omitted here.

It should be mentioned that the concept of the DAO implies that multi-sensor image fusion is generally not a “pure” image processing problem. Rather, researchers must pay more attention to the physical implications of images. We deem that the new concepts are more practical and able to help researchers in our area to extend their considerations in the research of image fusion.

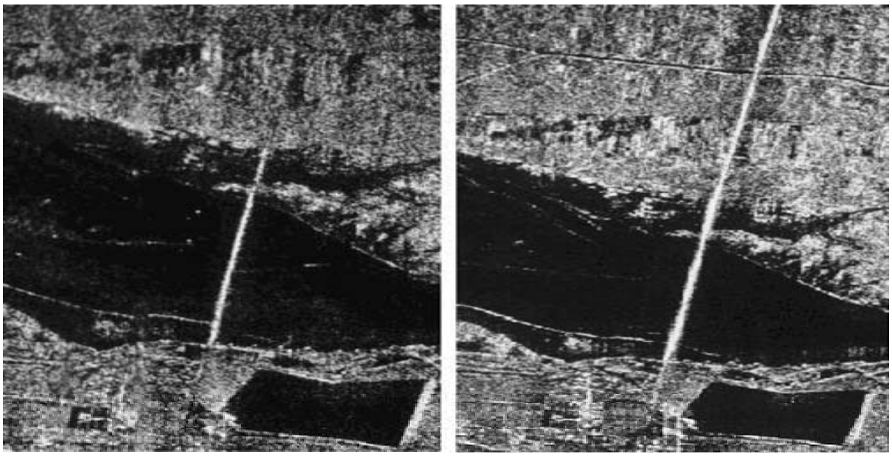
2. The object contained in the image

The primary difficulty faced by image fusion comes from the understanding of what the object or content is in the image. From the view point of image processing, researchers are accustomed to establish a direct correspondence between a set of pixels and an object in the image. Physically, however, some prejudice hides behind this understanding. People used to find and interpret objects from a clear optical picture. Actually, even if the picture is high-resolution, it is not rare the case that it cannot contain complete information that we want, e.g. if the object is masked by clouds, leaves, tents or even paint. Microwave and infrared sensors, on the other hand, may detect the objects’ attributes that possibly cannot be obtained by optical sensors. For example, The L-band SAR (Synthetic Aperture Radar) can find metal objects through masking of clouds, leaves, tents or paint. Additionally, the infrared imaging devices can reflect the temperature distribution of the object itself and its neighborhood. Usually, the optical image, the radar image, and the infrared image of the same object may be very different in appearance, pattern, and size. In such a case, so-called ‘pixel-level image fusion’, which aims to improve image resolution, will become lack of scientific ground. Actually, what we need is such a technique that can implement information fusion by making fully use of image information we obtained. In order to find and identify an object reliably, we not only need the object’s appearance but also the object’s various physical attributes. Fig.1 shows three image samples. Fig.1 (a) is a multi-spectral image of somewhere, which shows more detailed geographical characteristics. Fig.1 (b) is its infrared image of the same area as that of the image in Fig.1 (a), which displays the temperature distribution of the area. Fig.1 (c) is a piece of a SAR image, in which there are a few airplanes. We cannot see any airplane clearly in Fig.1 (c). However the ultra-intensive reflection pattern appeared in the image shows the existence of the metal objects. Furthermore, the reflection pattern can offer useful information for confirming the structural characteristics of airplanes.

（a） (b) (c)

Fig.1 Three image samples. (a) A multispectral image. (b) An infrared image of the same area as that of the image in (a). (c) A piece of SAR image.

In multi-spectral image processing, images at different wave band and different polarization mode of the same vegetation area are usually different. In fact, such difference exhibits the property of the vegetation, i.e. it contains the feature attributes for distinguishing the property of the vegetation. In order to extract and use these attributes, we should analyze every image and relate all attributes of all images together. But this by no means implies that we must ‘fuse’ these images together and form a single one. For such application, usually, it would be more desirable to create a few images or pseudo-colored geographic maps, each one of them exhibits a distribution of a special attribute of the vegetation area.



(a) (b)

Fig.2 Images of the same metal bridge imaged by different polarization SAR. （a）VV Polarization SAR image； （b）HH Polarization SAR image

Polarization SAR images are an important and typical example. The theory and practice of backscattering of electromagnetic wave show that the reflection will enhance if the orientation of an edge of the object is coincident with the polarization direction of the electromagnetic wave. Especially, reflection will become much intensive if the length of the reflection edge is closer to the electromagnetic wavelength. If a SAR can work in HH, VV, or VH polarization modes, the acquired images in different modes will be different in general, even if the object has not any change. Fig.2 shows two images of the same metal bridge imaged by a SAR in different polarization modes. It can be seen that the lengths of the bridge in two images are quite different. This difference just reflects the metal attribute of the bridge. Imaging property caused by polarization has evident manifestations not only for metal objects but also for vegetations, soil and so on, although the manifestations may be different. It means that SAR images and even optical images obtained by various polarizations are different for different earth's surface, i.e. these images contain the feature attributes of earth's surface. These results of observation and research have been widely used in agricultural monitoring, land and resource surveying, disaster evaluation and statistics [Rignot, Camps-Valls 2008, Aanæs 2008, Unsalan 2009, Hedman et al. 2010].

In conclusion of our observation, an image should be considered as a distribution of attributes of objects (DAO). A distribution (image) acquired from a sensor under a specified working mode describes certain particular attributes of the object. Different distributions (images) describe the object characteristics from different points of view, such as the visual appearance, material, the structure size and orientation, temperature, the behavior of reflection, radiation and scattering, the foreground and the background associated with objects, etc. The task of image fusion is to utilize all information completely so as to realize information fusion.

3. Key techniques involved in image fusion

From the above observation, it can be seen that the research and application of image fusion depend on a mass of image materials coming from various imaging sensors. It is unlikely possible to provide reliable results using technical methods with absence of knowledge accumulation. Accompanied with image acquiring, analyzing and knowledge accumulation, the development of related theory and techniques becomes indispensable.

(1) ***Multi-sensor image registration***

Image registration based on the mutual correlation is suitable only to the case when two images to be aligned have the same gray feature. In medical image processing, in order to realize the registration operation between a CT (Computerized Tomography) image and an MRI (Magnetic Resonance Imaging) image, researches have developed a multi-modal image registration technique [Maintz 1998, Wong 2008]. The underlying theoretical basis is statistically computing the degree of coincidence of the variation of gray distributions of two images' pixels. Such a degree of coincidence can be described by Kullback-Leibler divergence, mutual information, or mutual variance. Multimodal image registration can be used to process multi-sensor remote sensing image registration [Wang 2002, Shi 2004, Woo et al. 2009, Li et al. 2009]. However, the development of the related theory and technique remains expected.

Inevitably, there are geometric distortions among multi-sensor remote sensing images, which make direct large-scale image registration non-practical. A technical method for the correction of geometric distortions is needed first. A typical geometric correction method is to determine two sets of feature points respectively in two images. The two sets of feature points are correctly corresponding to each other geometrically and they are used as reference points for geometric correction of one image to another. Traditional feature points are crossing points, corner points, etc. In practice, however, these points are difficult to be determined at pixel level for lack of a determinate criterion. In such a case, the correctness of selected feature points is not guaranteed. The concept of generalized feature points (GFD) [Zou 2002] provides more reasonable ground for determining the reference points of geometric correction.

The primary idea of the GFD concept is as follows [Zou 2002]. Let A and B be two images to be registered. Select a small sub-image in A and it is denoted by SA1. The small image SA1 should be complicated enough, e.g. it contains crossing points, corners, endpoints, etc. We search for a small sub-image in B, denoted by SB1, so that SB1 and SA1 can be registered to each other exactly. This is possible by using the multimodal image registration technique. If this is realizable, an arbitrary point in SA1 has its counterpart in SB1, and such a pair can be used to serve as the feature points and called as the GFD. Typically, the centers of SA1 and SB1 can be selected as a pair of GFD. Similarly, we can select and determine more pairs of GFD in A and in B.

For multi-sensor image fusion, we need not only to register/align the related images, but also to exhibit the different attributes of different images. This implies that, in certain areas of images, the patterns would be different from one image to the others. These different patterns cannot be registered each other. Rather, we must conserve these differences respectively. Clearly, we need such a registration technique that has the ability of distinguishing the object’s regions with non-object’s regions of images. Registration should be carried out only on the basis of non-object’s regions. In application, it is not really needed to determine the non-object’s regions exactly. In fact, for the aim of geometric correction we need only to determine three or more pairs of GFD located in the non-object’s regions. This is not a difficult task in usual case. However, detection of different patterns in different images by computer will be by no means an easy task for multi-model images. In fact, how to define “the different patterns” and distinguish it with other part of the image will be problematic. No exact theory and method have been reported in the literature.

From the theory and experimental results of multi-modal image registration, we can see that the reported techniques in the literature suit for the situation that the geometric distribution of edges of the two images to be aligned is generally coincident while the gray attributes of corresponding regions of the images may be not consistent. For multi-sensor image registration application, these conditions may be difficult to be satisfied. Actually, in order to realize registration of Fig.1 (a) with

Fig.1 (b), we need re-examine the existing theory and technique of multi-modal image registration because the lack of theory and method to verify and evaluate the precision of registration. Furthermore, it remains to be expected to investigate the registration problems for 3-D images and large size images formed from piecing together several images with distortion in geometry and in gray level. The handling of these problems would be challenging for researchers in this area.

(2) ***Segmentation of the commonly interesting region***

On the basis of image registration, the second preprocessing step of implementation of image fusion is segmentation of the commonly interesting region in each related image. This work is naturally application-oriented. The interesting region may be the object itself. Also it may be the region containing the objects or the circumstance where objects are located. Segmentation of regions manually is not so realistic for massive image processing. In the literature, we can find a number of reports for image segmentation [Pham 2004, Artaechevarria et al. 2009, Wan et al. 2009]. The most useful segmentation techniques can be established on the basis of the homogeneity of attributes in the region [Zou 2002]. The homogeneity generally used for image segmentation includes the homogeneity of gray level attribute; of statistical features of textures; of texture patterns; and of velocities of objects (for video image segmentation). A typical method is to compute the features of images in a specified small window. The features typically are: parameters of the probabilistic or deterministic image model, parameters of the images in the transform domain (Fourier Transform, DCT, Gabor Transform, Wavelet Transform, Curvelet Transform or Multiresolution Geometric Analysis, Singular Value Decomposition, Gray Histogram, etc.). These features can be directly used in image segmentation. However, a significant error of segmentation will appear in the neighborhood of the boundary of the region by this method. The active contour method has been successfully used for segmentation of images with homogeneity in gray level and it is very suitable to the task of segmentation and extraction of ocean coasts, lake surfaces, river channels and suchlike [Terzopoulious 1986, Kass et al. 1987]. There are lots of reports about image segmentation with large scale texture. But more general and efficient techniques are seeking yet [Zhang 2002, Yang 2004].

Segmentation technique based on the feature homogeneity can only deal with very limited segmentation problems. Image fusion possibly requires more additional regional assignment techniques. For example, in order to evaluate the disaster, we need not only to determine the disaster-influenced region but also the potentially and secondarily disaster-influenced areas. At the same time, we usually need to evaluate not only the capital disaster but the secondary disaster and the intergrowth disaster. When all of information should be exhibited on the basis of a complete set of images acquired from the past to the current, the image segmentation technique required for image fusion application will be in a composite form. In such a case, image segmentation and area partition need a data support from the large-area monitoring system. From the system framework of data fusion technique, such a level of inter-hierarchy fusion is short of investigation. Planning, controlling and management of the urbanization process are important application of image fusion. It needs more detailed regional assignment and segmentation techniques to provide various statistics of information, such as constructing/constructed urban areas, planning regions, geography and geological distribution, and the dynamic process of urban evolution, etc. However, from the view point of image processing, how to define the segmentation region lacks of proper standard or criterion because the urban architecture usually forms irregular image patterns [Akçay 2007, 2008, Hedman et al. 2010].

In practice of multi-sensor image fusion, the objects may exhibit divergent patterns and the boundaries of the objects versus its background may be not clearly distinguishable. This is the real case especially for radar and infrared images. In such a case, the conventional segmentation technique would produce a profile that may be different greatly from the real profile of the objects. A plausible way to avoid this difficulty is to produce a compact object's region (COR) which satisfies a few conditions as follows. 1) The COR should be large enough that can contain the complete object's pattern, visible and possibly invisible. 2) It is a common region for all acquired images from different sensors. 3) It is a compact (small enough) region that satisfies the conditions 1) and 2). To this end, the segmentation technique should be in a composite form that integrates several criteria and should be jointly carried out simultaneously to all images. The COR is an effective basis for extracting and describing the feature attributes of objects.

(3) ***Extraction and association of feature attributes of objects***

Image processing for application of image fusion is asked to offer warranted information for the final image interpretation, and therefore extracting and describing the feature attributes of objects become very important. The attributes of an image region generally mean the features of the texture of the region. We can extract various features of textures by means of mathematic modeling and transform techniques of images [Romdhani 2007]. Since an image is understood as a DAO, a basic problem should be investigated: what kind of feature attributes can fully reflect the physical property of a specified kind of objects and what kind of mathematic models is the most suitable for extracting the feature attributes. For example, the utilization of the polarization attributes has comprehensive significance. How to jointly process various texture features of polarized SAR images needs deeper theoretical study and experimental investigation.

Extraction of objects from an image is a basic operation in image processing. Usually, a segmentation technique is required for extracting objects contained in an image. Typically, the profiles of objects are extracted. But this is not always feasible, when the boundary of the object is not distinguishable with the background. In military image fusion, a kind of particular artificial objects, such as airports, roads, large buckets, bridges, etc, are highly concerned [Mayer 1999, Akçay 2007, 2008]. Usually, we cannot extract these objects directly from a large size image because in such a way we will waste too much computation resource. The grid searching scheme has the similar difficulty. At the first step, we can assign a particular object region that contains the objects. This can be done on the basis of history knowledge obtained from remote sensing for most objects. For emerged unknown objects and moving objects, such as airplanes, vehicles, tanks, warships, etc, a seeking algorithm would be needed. In any case, if a compact region that contains objects can be assigned before seeking, the computational burden of extraction operation will decrease drastically.

In the development of the multimedia content description standard, MPEG7, researchers have proposed a series of technical methods for extracting and describing the content in images [Manjunath 2002]. The research results for MPEG7 inherently can be utilized as a good reference for researchers in our community. Take care that, however, implementation of image fusion requires much more precise results of extraction and description. A systemic research for the theory, method, and criteria remains in expectation.

To describe regular artificial objects, such as airports, roads, large buckets, bridges, etc., there exist a number of methods that can be utilized singly or compositely. The mathematic transform methods (Fourier Transform, DCT, Gabor Transform, Wavelet Transform, Curvelet Transform or Multi-resolution Geometric Analysis, Singular Value Decomposition, Gray Histogram, etc.) are commonly used for extracting the numerical features of objects. Complex moments are a kind of numerical features with invariability. These features have been used successfully for pattern identification [Abu-Mostafa 1985, Jin and Davis 2005, Ruvimbo 2009]. Hough transform is a feature extraction technique used for finding imperfect instances of objects within a certain class of shapes [Shapiro 2001]. The result of Hough transform can offer features of positions and arbitrary shapes including lines, squeres, circles or ellipses.

Essentially, the description of an object by a set of numerical features implies a kind of simplified expresion of that object. A question is: what kind of description is the most efficient? The question closely relates to several important scientific problems such as data complexity, modeling, compression, and encoding. The research on this problem for images is highly expected.

There is lack of general methods for describing movable objects, such as planes, vehicles, tanks, etc. These metal objects have strong reflecting points or reflecting lines accompanied by a particular reflection pattern in SAR images. Also they may have certain temperature distribution pattern in infrared images. Such a pattern will form a contrast with the background. Moreover, the movability of objects is a useful attribute. SAR will produce an imaging abnormality for moving objects. This implies a possibility of detecting moving objects. Furthermore, it is an effective method for images acquired at different time to obtain information of moving objects through moving detection. Care is needed that this method relies on the precision of image registration. The description method for attributes of moving metal objects now is not at hand. For SAR sensors, a possible method is to establish a database of reflection modes of various typical objects. Unluckily, the reflection patterns of SAR objects are greatly diverse. The reflection patterns depend on many factors, such as radar working frequency, transceiver polarization arrangement, object's external structure, the spatial position and attitude of the object relative to the radar, etc. This makes the database of reflection patterns too large to be established. On the basis of reflection patterns, a possible way for establishing the radar object database is to find a set of numerical features with some invariability that can describe the reflection patterns. But it needs further theoretical and experimental study.

Association of feature attributes of objects has its special meaning for multi-sensor image fusion. The all attributes of an object obtained from all acquired multi-sensor images are associated to each other. This will offer a sound base for describing the objects so that the primary idea of image fusion can be implemented.

To this end, an *association database* should be established. The association database proposed here is a database method which automatically associates the complete set of numerical features coming from different sensors relating to the same object or region. In the association database, a description may contain several terms. Each term consists of a few parameters that reflect the attributes of the object from all possible view angles. This is just the way of "blind men apperceive a big elephant". For a special kind of objects, we should study and investigate the most essential features and the most suitable sensors. The database permits users to access, use and display each single image or combine all related images together flexibly, and to implement the synthesis and fusion algorithms conveniently. Such an association database must be application-oriented. For a specified application, a special data protocol, *the object's attribute-feature protocol* (OAFP), should be established at the first step.

Imaging a practical situation, if a few multi-sensor images of a same interesting area have been acquired, a primary question should be answered is that how to associate or relate the contents or objects of every image together? A general theory and method to answer this question would be very difficult to be made out. For a specified and a limited application, *the model-based attribute association* of objects is possible. For military image fusion application, for example, the objects to be concerned are limited. The available imaging apparatuses are also limited. Through a massive experiment investigation, the set of feature-attributes of objects under different imaging conditions can be established. Such a set can be viewed as a model for describing the object. This forms a basis for the model-based attribute association of objects and therefore the basis of the OAFP. Once such an OAFP is architected, it will offer a unified work frame in the research and application of image fusion.

(4) ***Information fusion and automatic image interpretation***

This is the step of producing results of image fusion. On the basis of the established association database, the implementation of image fusion from the newly obtained images will become reasonable. The published information fusion theory and technique mainly concern the application of data fusion and fusion decision [Hall 2001]. The typical applications are concentrated to detection, tracking, and identification of multiple moving targets; battlefield intelligence, surveillance, and situation assessment; and integration and decision in Command-Control-Communication & Intelligence (C^3I) systems; and a few non-military fields [Bar et al. 2010]. Image fusion has its extraordinary character which stems from the trait of images: richness and diversity of the content, and visibility and intuitiveness. In fact, the so-called "content" of an image may be divergently interpreted in usual case. In application, researchers must confine themselves to search for a specified set of attributes, each one of which can be described by a small set of numerical features. This would be realizable in virtue of the model-based attribute association of objects and the special data protocol OAFP mentioned previously.

Automatic image interpretation is highly desirable for application. Image interpretation closely relates to pattern recognition and image understanding. The research on these topics has got fruitful results in the recent several ten years. In the multi-sensor imaging circumstance, the input entries would be multiple. Once the operation of object association can be carried out on the basis of the OAFP, the development of the pattern recognition and image interpretation algorithms can be schemed on the brand-new basis. Since the objects we concerned can be described through multiple features, and each feature can be described by several numerical values, a weighted-composite feature-matching method would work well. The further advanced recognition and interpretation techniques may include [Dai and Khorram 1999, Yun 2004, Dong et al. 2004, Huang and Jing 2007, Wang et al. 2007, Theodoridis, Koutroumbas 2009, Howson and Urbach 2005, Bolstad 2007, Carlin and Louis 2008, Bolstad 2010]:

Hybrid numerical - morphological identification method
Bayesian statistic based method
Artificial neural network based identification method
Mathematic manifold based identification method

With the hybrid and composite inputs, the mentioned advanced methods and related theory should be re-examined. Nevertheless, the available data are now richer than ever, the correctness of object's identification would be warranted with relative ease.

When the result of image fusion is displayed in a dynamic video form or in the form of contrasted image series or geographic map series, the impact to our eye and thought would be extraordinary. This implies that the mental factor and experience should be considered in fusion decision and image interpretation. In fact, a complete automatic image interpretation for practical scale images will ask a huge volume of computational resource. In practical circumstance, with the help of an experienced image interpreter, the man-machinery interactive image interpretation system would be much efficient. The interpreter can help to restrict the size of the concerned areas; to tag a part of intuitively discriminable objects; to correct some of errors in interpretation; and more. Clearly, in addition to the automatic image identification and fusion ability, a practical and efficient image fusion system should allow the control inputs from interpreters.

4. Conclusions

Since an image can be understood as a DAO, multi-sensor image fusion is naturally the matter of attribute composition or information fusion. With this view point, researchers must pay more attention to the physical implications of images. In the multi-sensor circumstance, the patterns of an identical object in different images may be significantly different from each other. The differences among patterns contain the feature attributes of the object. In this paper, a few preprocessing techniques, such as multi-sensor image registration, segmentation of the commonly interesting region, and extraction of feature attributes of objects, are discussed with considerations relating to image fusion. With the concept of the DAO, feature-level image fusion turns out to be association of attributes of objects. The natural reasoning of the consideration leads to the proposals for the association database and the OAFP, which can be established on the basis of the model-based attribute association of objects. By means of the established association database, the implementation of image fusion from the newly obtained images will become reasonable.

From the proposed concepts, the popular definition for image fusion, “The combination of two or more different images to form a new image by using a certain algorithm”, could not be the essential. “To form a new image” can be substituted by “to produce a set of associated attribute descriptions of objects”. The display and the related techniques are important, but not limited to a single fused image.

The DAO is a new concept. It brings about a series of problems that should be considered in multi-sensor image fusion. In this paper, only a few commonly encountered problems have been mentioned. However, the range of the related research work is large, multifold, and challenging. The multi-sensor image fusion technique and the related theoretical concepts need continue efforts of many researchers.

Notes on contributors

Zou Mou-yan, Research Professor and retired from the Institute of Electronics, Chinese Academy of Science; Professor of the Graduate School, Chinese Academy of Science; Chief Executive Director of the Beijing TransMicrowave Technology Corp.. He has published more than 70 academic papers and is the author of the book “Deconvolution and Signal Recovery”. The recent research interests include image processing and commercial microwave radar techniques with application to the road traffic control and safety.

Liu Yan, Lecturer of the Graduate School, Chinese Academy of Science. She Received her Doctoral Degree in 2007 from the Institute of Electronics, Chinese Academy of Science. The recent research interests include image processing and video techniques with applications.

References

Aanæs, H., Sveinsson, J. R., and Nielsen, A. A. et al., 2008, Model-Based Satellite Image Fusion, *IEEE Transactions on Geoscience and Remote Sensing*, 46(5)，1336-1346.

Abu-Mostafa, Y. and Psaltis, D., 1985, Image normalization by complex moments. *IEEE Trans Pattern Analysis Machine Intelligence*, PAMI-7:46–55.

Akçay, H. G. and Aksoy, S., 2008, Automatic Detection of Geospatial Objects Using Multiple Hierarchical Segmentations, *IEEE Transactions on Geoscience and Remote Sensing*, 46(7)，2097-2111.

Akçay, H. G. and Aksoy, S., 2007，Automated Detection of Objects Using Multiple Hierarchical Segmentations, *Proc. of IEEE International Geoscience and Remote Sensing Symposium* (IGARSS’07) 1468-1471.

Artaechevarria X., Munoz-Barrutia A., Ortiz-de-Solorzano C., 2009, Combination Strategies in Multi-Atlas Image Segmentation: Application to Brain MR Data, *IEEE Transactions on Medical Imaging*, 28(8), 1266-1277.

Bai Ying-Wen, Shen Li-Sih, Li Zong-Han, 2010, Design and implementation of an embedded home surveillance system by use of multiple ultrasonic sensors, *IEEE Transactions on Consumer Electronics*, 56(1), 119-124.

Bar D.E., Wolowelsky K., Swirski Y.，et al., 2010, Target Detection and Verification via Airborne Hyperspectral and High-Resolution Imagery Processing and Fusion, *IEEE Sensors Journal*,10(3), 707-711.

Blum, R.S.; Liu, Z. 2006, *Multi-Sensor Image Fusion and Its Applications*; special series on Signal Processing and Communications; CRC Press: Boca Raton, FL, USA.

Bolstad, William M., 2007, Introduction to Bayesian Statistics: Second Edition, John Wiley.

Bolstad, William M. (2010) Understanding Computational Bayesian Statistics, John Wiley.

Bovolo F., Bruzzone L., Capobianco L., et al., 2010, Analysis of the Effects of Pansharpening in Change Detection on VHR Images, *IEEE Geoscience and Remote Sensing Letters*, 7(1), 53-57.

Brunner D., Lemoine G., Bruzzone L., 2010, Earthquake Damage Assessment of Buildings Using VHR Optical and SAR Imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 48(5), 2403-2420.

Calhoun V.D., Adali T., 2009, Feature-Based Fusion of Medical Imaging Data, *IEEE Transactions on Information Technology in Biomedicine*, 13(5), 711-720.

Camps-Valls, G. Gómez-Chova, L. and Muñoz-Marí, J. et al., 2008, Kernel-Based Framework for Multitemporal and Multisource Remote Sensing Data Classification and Change Detection, *IEEE Transactions on Geoscience and Remote Sensing,* 46(6)，1822-1835.

Carlin, Bradley P. and Louis, Thomas A. (2008). Bayesian Methods for Data Analysis, Third Edition. Boca Raton, FL: Chapman and Hall/CRC.

Cartmill J., Wachowski N., Azimi-Sadjadi M.R., 2009, Buried Underwater Object Classification Using a Collaborative Multiaspect Classifier, IEEE Journal of Oceanic Engineering, 34(1), 32-44.

Celik T., Ma Kai-Kuang, 2010, Unsupervised Change Detection for Satellite Images Using Dual-Tree Complex Wavelet Transform, *IEEE Transactions on Geoscience and Remote Sensing*, 48(3), 1199-1210.

Chen Shaohui, Zhang Renhua, Su Hongbo, et al, 2010, SAR and Multispectral Image Fusion Using Generalized IHS Transform Based on à Trous Wavelet and EMD Decompositions, *IEEE Sensors Journal*, 10(3), 737-745.

Copeland A. D., Mangoubi R. S., Desai M. N., et al., 2010, Spatio-Temporal Data Fusion for 3D+T Image Reconstruction in Cerebral Angiography, *IEEE Transactions on Medical Imaging*, 29(6), 1238-1251.

Correa Nicolle M., Adali Tulay, Li, Yi-Ou, et al., 2010,Canonical Correlation Analysis for Data Fusion and Group Inferences, *IEEE Signal Processing Magazine,* 27(4), 39-50.

Cvejic N., Seppanen T., Godsill S.J., 2009, A Nonreference Image Fusion Metric Based on the Regional Importance Measure, *IEEE Journal of Selected Topics in Signal Processing*, 3(2), 212-221.

Dai, X.; Khorram, S. Data fusion using artificial neural networks: a case study on multitemporal change analysis. *Comput. Environ. Urban Syst*. 1999, *23,* 19–31.

Dong. J.; Yang, X.; Clinton, N.; Wang, N. 2004, An artificial neural network model for estimating crop yields using remotely sensed information. *Int. J. Remote Sens.*, *25*, 1723–1732.

Dong Jiang, Zhuang Dafang, Huang Yaohuan and Fu Jingying，2009，Advances in Multi-Sensor Data Fusion: Algorithms and Applications，*Sensors, 9,* 7771-7784.

Flusser, Jan. Sroubek, F. Zitova, B., 2007，Image Fusion Principles Methods and Applications, Tutorial of *the 15th European Signal Processing Conference* (EUSIPCO’2007).

Gamba, P. Hellwich, O. and Lombardo, P. 2005，Fusion of Remotely Sensed Data Over Urban Areas, *Information Fusion*, (6)，189-192.

Goshtasby, A. A. and Nikolov, S., 2007，Image fusion: Advances in The State of The Art, *Information Fusion*, (8)，114-118.

Gundimada S., Asari V.K., 2009, Facial Recognition Using Multisensor Images Based on Localized Kernel Eigen Spaces, *IEEE Transactions on Image Processing*, 18(6), 1314-1312.

Hall, D.L. and Llinas, J., 1997, An Introduction to Multisensor Data Fusion, *Proc. IEEE*, 85 (1)，6–23.

Hall, D. L. and Llinas, J., ed., 2001, *Handbook of multisensor data fusion*, CRC Press, New York.

Hedman K., Stilla U., Lisini G., et al., 2010, Road Network Extraction in VHR SAR Images of Urban and Suburban Areas by Means of Class-Aided Feature-Level Fusion, *IEEE Transactions on Geoscience and Remote Sensing*, 48(3), 1294-1296.

Hizem W., Allano L., Mellakh A., et al., 2009，Face recognition from synchronised visible and near-infrared images, *IET Signal Processing*, 3(4), 282-288.

Howson, Colin and Urbach, Peter, 2005, *Scientific Reasoning: the Bayesian Approach* (3rd ed.). Open Court Publishing Company.

Huang, W.; Jing, Z. Multi-focus image fusion using pulse coupled neural network.

Pattern Recognit. Lett. 2007, 28, 1123–1132.
Jeffrey K. A., Otto L. L., Kozioziemski, Bernard J., et al., 2009, Refraction-enhanced x-ray radiography for inertial confinement fusion and laser-produced plasma applications, *Journal of Applied Physics*, 105(11), 113112-113112-8.
Jin, X.Y.; Davis, C.H. An integrated system for automatic road mapping from high-resolution multi-spectral satellite imagery by information fusion. *Inf. Fusion* 2005, *6*, 257–273.
Joshi M., Jalobeanu A., 2010, MAP Estimation for Multiresolution Fusion in Remotely Sensed Images Using an IGMRF Prior Model, *IEEE Transactions on Geoscience and Remote Sensing*, 48(3), 1245-1255.
Jun S., Xiaoling Z., Jianyu Y., et al., 2009, Radix- Resolution-Fusion for LASAR via Orthogonal Complement Decomposition, *IEEE Geoscience and Remote Sensing Letters*, 6(1) , 147-151.
Kalka N.D.; Jinyu Zuo, Schmid N.A.,et al., 2010, Estimating and Fusing Quality Factors for Iris Biometric Images, *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 40(3), 509-524.
Karantzalos K., Paragios N., 2010, Large-Scale Building Reconstruction Through Information Fusion and 3-D Priors, *IEEE Transactions on Geoscience and Remote Sensing*,48(5), 2283-2296.
Kass M., Witkin A., Terzopoulos D., 1987, Snakes: Active Contour Models, International Journal Comp. Vis., 1(4):321-331.
Kumar M. , Dass S., 2009, A Total Variation-Based Algorithm for Pixel-Level Image Fusion, *IEEE Transactions on Image Processing*, 18(9), 2137-2143.
Kumar, R. Anandan, P. Irani, M., etc., 1995，Representation of Scenes from Collections of Images, *Proc. of IEEE Workshop on Representations of Visual Scenes*, 10-17.
Li Zhenhua, Leung H., 2009, Fusion of Multispectral and Panchromatic Images Using a Restoration-Based Method, *IEEE Transactions on Geoscience and Remote Sensing*, 47(5) , 1482-1491.
Li Qiaoliang, Wang Guoyou, Liu Jianguo, et al., 2009, Robust Scale-Invariant Feature Matching for Remote Sensing Image Registration, *IEEE Geoscience and Remote Sensing Letters*, 6(2), 287-291.
Licciardi G., Pacifici F., Tuia D., et al., 2009, Decision Fusion for the Classification of Hyperspectral Data: Outcome of the 2008 GRS-S Data Fusion Contest, *IEEE Transactions on Geoscience and Remote Sensing*, 47(11), 3857-3865.
Looney D., Mandic D.P., 2009, Multiscale Image Fusion Using Complex Extensions of EMD, *IEEE Transactions on Signal Processing*, 57(4), 1626-1630.
Ma Hongchao, Yang Yun, 2009, Two Specific Multiple-Level-Set Models for High-Resolution Remote-Sensing Image Classification, *IEEE Geoscience and Remote Sensing Letters*, 6(3), 558-561.
Masini A., Corsini G., Diani M., et al., 2009, Analysis of Multiresolution-Based Fusion Strategies for a Dual Infrared System, *IEEE Transactions on Intelligent Transportation Systems* , 10(4), 688-694.

Maintz, J.B.A., Viergever, M.A., 1998, A Survey of Medical Image Registration, *Medical Image Analysis*, 2(1)，1–36.
Manjunath, B.S., Salembier, P. and Sikora, T. Ed., 2002, *Introduction to MPEG-7: Multimedia Content Description Interface*. Wiley & Sons.
Molch K., 2010, Multisensor and Multitemporal Fusion of VHR Satellite Imagery Based on KIM, *IEEE Geoscience and Remote Sensing Letters*, 7(1), 48-52.
Monwar M.M.; Gavrilova M.L., 2009, Multimodal Biometric System Using Rank-Level Fusion Approach, *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 39(4), 867-878.
Moser G., Serpico S.B., 2009,Unsupervised Change Detection From Multichannel SAR Data by Markovian Data Fusion, *IEEE Transactions on Geoscience and Remote Sensing*, 47(7), 2114-2128.
Orchard J., Mann R., 2010,Registering a MultiSensor Ensemble of Images, *IEEE Transactions on Image Processing*, 19(5), 1236-1247.
Pham D.L., Xu Chenyang, and Prince J.L. 2004, A survey of current methods in medical image segmentation. http://www.rfai.li.univ-tour.fr/rousselle/docum/pdf/p124r.pdf.
Pastina D., Spina C., 2009, Multi-feature based automatic recognition of ship targets in ISAR, *IET Radar, Sonar & Navigation*, 3(4),406-423.
Petillot I., Trouve E., Bolon P.,et al., 2010, Radar-Coding and Geocoding Lookup Tables for the Fusion of GIS and SAR Data in Mountain Areas, *IEEE Geoscience and Remote Sensing Letters*,7(2), 309-313.
Poh N., Kittler J., Bourlai T., 2010, Quality-Based Score Normalization With Device Qualitative Information for Multimodal Biometric Fusion, *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 40(30), 539-554.
Pohl, C., Genderen, J. L., 1998, Multisensor Image Fusion in Remote Sensing: Concepts, Methods and Applications, *International Journal of Remote Sensing*, 9(5)，823-854.
Qian M. ,　Aguilar　M., Zachery K.N., et al., 2009, Decision-Level Fusion of EEG and Pupil Features for Single-Trial Visual Detection Analysis, *IEEE Transactions on Biomedical Engineering*, 56(7), 1929-1937.
Rignot, E. and Chellappa, R., Segmentation of Polarimetric Synthetic Aperture Radar Data, *IEEE Trans. on Image Processing*, 1(3)，281 – 300.
Romdhani, S. and Vetter, T., 2007, 3D Probabilistic Feature Point Model for Object Detection and Recognition, *Proc. of IEEE Conference on Computer Vision and Pattern Recognition* (CVPR '07) 1-8.
Ruvimbo, G.; Philippe, D.; Morgan, D. 2009, Object-oriented change detection for the city of Harare, Zimbabwe. *Exp. Syst. Appl. 36*, 571–588.
Shah V.P., Younan N.H., Durbha S.S., et al., 2010, Feature Identification via a Combined ICA–Wavelet Method for Image Information Mining, *IEEE Geoscience and Remote Sensing Letters*, 7(1), 18-22.
Shapiro, L. and Stockman, G., 2001, *Computer Vision*, Prentice-Hall, Inc..
Shi Yonggang, 2004，Research of Generalized Distance Measure and Multi-modal

Image Registration, PhD Thesis, Institute of Electronics, Chinese Academy of Sciences (in Chinese).

Simone, G. Farina, A. and Morabito, F.C. et al., 2002，Image Fusion Techniques for Remote Sensing Applications, Information Fusion, (3)，3-15.

Smith, M.I.; Heather, J.P. 2005, Review of image fusion technology in 2005. In *Proceedings of Defense and Security Symposium*, Orlando, FL, USA.

Terzopoulios D., 1986, On matching deformable models to images, Technical Report 60, Schlumberger Palo Alto Research,

Theodoridis, Sergios, Koutroumbas, Konstantinos, 2009, "Pattern Recognition", 4th Edition, Academic Press, ISBN: 978-1-59749-272-0.

Thomas, C., Ranchin, T., and Wald, L.,et al., 2008, Synthesis of Multispectral Images to High Spatial Resolution: A Critical Review of Fusion Methods Based on Remote Sensing Physics, IEEE Transactions on Geoscience and Remote Sensing, 46(5)，1301-1312.

Udelhoven T., van der Linden S., Waske B.， et al., 2009, Hypertemporal Classification of Large Areas Using Decision Fusion, *IEEE Geoscience and Remote Sensing Letters*, 6(3),592-596.

Unsalan C., 2009, Statistical, Structural, Hybrid, and Graph Theoretical Features to Measure Land Development, *IEEE Geoscience and Remote Sensing Letters*, 6(1), 72-76.

Vijayaraj, V.; Younan, N.; O‟Hara, C. Concepts of image fusion in remote sensing applications. In *Proceedings of IEEE International Conference on Geoscience and Remote Sensing Symposium*, Denver, CO, USA, July 31–August 4, 2006; pp. 3798–3801.

Wan Tao, Canagarajah N., Achim, A., 2009, Segmentation-Driven Image Fusion Based on Alpha-Stable Modeling of Wavelet Coefficients, *IEEE Transactions on Multimedia*,1 1(4), 624-633.

Wang Dongfeng, 2002，A Study on Registration of Multimodal and Huge-size images, PhD Thesis, Institute of Electronics, Chinese Academy of Sciences (in Chinese).

Wang, R.; Bu, F.L.; Jin, H.; Li, L.H. 2007, A feature-level image fusion algorithm based on neural networks. *Bioinf. Biomed. Eng.*, *7*, 821–824.

Wang Z.F., Han Q., Li Q., et al., 2009, Complex common vector for multimodal biometric recognition, *Electronics Letters*, 45(10),495-496.

Wong, A. and Bishop, W., 2008，Efficient Least Squares Fusion of MRI and CT Images Using A Phase Congruency Model, *Pattern Recognition*, (29)，173–180.

Woo Sang-Keun, Kim Kyeong Min, Lee Tae Sup, et al., 2009, Registration Method for the Detection of Tumors in Lung and Liver Using Multimodal Small Animal Imaging, *IEEE Transactions on Nuclear Science*, 56(3), 1454-1458.

Wu Shunguang, Decker S., Chang Peng, et al., 2009,Collision Sensing by Stereo Vision and Radar Sensor Fusion, *IEEE Transactions on Intelligent Transportation Systems*, 10(4), 606-614.

Xie Shufu, Shan Shiguang, Chen Xilin, et al., 2010, Fusing Local Patterns of Gabor

Magnitude and Phase for Face Recognition, *IEEE Transactions on Image Processing*, 19(5), 1349-1361.

Yang Bin, Shutao Li, 2010, Multifocus Image Fusion and Restoration With Sparse Representation, *IEEE Transactions on Instrumentation and Measurement*, 59(4),884-892.

Yang Hongbo, 2004，Research of Unsupervised Texture Segmentation, PhD Thesis, Institute of Electronics, Chinese Academy of Sciences (in Chinese).

Yun, Z. 2004, Understanding image fusion. *Photogram. Eng. Remote Sens.*, *6*, 657–661.

Zhang Fan, Wu Xiaolin, Yang Xiaokang, et al., 2009, Robust Color Demosaicking With Adaptation to Varying Spectral Correlations, *IEEE Transactions on Image Processing*, 18(12),2706-2717.

Zhang Yifan, De Backer S., Scheunders P., 2009, Noise-Resistant Wavelet-Based Bayesian Fusion of Multispectral and Hyperspectral Images, *IEEE Transactions on Geoscience and Remote Sensing*, 47(11) , 3834-3843.

Zhang Lifei, 2002，Active Contour Technique Based on Minimum Description Length Principle, PhD Thesis, Institute of Electronics, Chinese Academy of Sciences (in Chinese).

Zou Mouyan, Wang Dongfeng, Zhang Lifei, 2002，“The Evolution and Key Problems of Remote Sensing Image Segmentation and Registration Techniques”, *Proc. of the eighth Symposium on the Remote Sensing, Telemetry and Remote Control* (in Chinese).

Zou Mouyan and Liu Yan, 2009, “Multi-Sensor Image Fusion: Difficulties and Key Techniques”, *Proc. of the 2nd International Congress on Image and Signal Processing*, CISP’09, Oct, TianJin, China.