ISSN : 2165-4069(Online)
ISSN : 2165-4050(Print)

IJARAI

International Journal of
Advanced Research in Artificial Intelligence

Volume 2 Issue 1

www.ijarai.thesai.org

A Publication of
The Science and Information Organization

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN ARTIFICIAL INTELLIGENCE

THE SCIENCE AND INFORMATION ORGANIZATION

www.thesai.org | info@thesai.org












Editorial Preface

From the Desk of Managing Editor…

"The question of whether computers can think is like the question of whether submarines can swim." — Edsger W. Dijkstra, the quote explains the power of Artificial Intelligence in computers with the changing landscape. The renaissance stimulated by the field of Artificial Intelligence is generating multiple formats and channels of creativity and innovation.

This journal is a special track on Artificial Intelligence by The Science and Information Organization and aims to be a leading forum for engineers, researchers and practitioners throughout the world.

The journal reports results achieved; proposals for new ways of looking at AI problems and include demonstrations of effectiveness. Papers describing existing technologies or algorithms integrating multiple systems are welcomed. IJARAI also invites papers on real life applications, which should describe the current scenarios, proposed solution, emphasize its novelty, and present an in-depth evaluation of the AI techniques being exploited. IJARAI focusses on quality and relevance in its publications.

In addition, IJARAI recognizes the importance of international influences on Artificial Intelligence and seeks international input in all aspects of the journal, including content, authorship of papers, readership, paper reviewers, and Editorial Board membership.

The success of authors and the journal is interdependent. While the Journal is in its initial phase, it is not only the Editor whose work is crucial to producing the journal. The editorial board members , the peer reviewers, scholars around the world who assess submissions, students, and institutions who generously give their expertise in factors small and large— their constant encouragement has helped a lot in the progress of the journal and shall help in future to earn credibility amongst all the reader members.

I add a personal thanks to the whole team that has catalysed so much, and I wish everyone who has been connected with the Journal the very best for the future.

Thank you for Sharing Wisdom!

Managing Editor
IJARAI
Volume 2 Issue 1 January 2013
ISSN: 2165-4069(Online)
ISSN: 2165-4050(Print)

Editorial Board

Reviewer Board Members

CONTENTS

CBR in the service of accident cases evaluating

Application to the safety of rail-road transport systems.

Lassaâd Mejri
IFSTAAR - French Institute of Science and
Technology of Transport, Planning and Networks
Faculty of Science of Bizerte
University of Carthage
Tunisia

Sofian Madi, Henda Ben Ghézala
RIADI-GDL Laboratory,
National School of Computer Science
University of Manouba
Tunisia

***Abstract*—This paper introduces a research aiming at the development of a decision support system concerning the approval of automated railway transportation systems. The objective is to implement a valuation method for the degree of compliance of the automated transportation system in-group of safety standards by the analysis of the scenarios of accident. To reach this target, we envisaged an approach Rex (Return of experience) who draws the lessons of accidents / incidents lived and/or imagined by the experts of the analysis of security in the IFSTAAR. Our approach consists in offering a decision support in the side of the experts of the certification based on a reuse of the scenarios of accidents already validated historically on other approved transportation systems. This approach Rex is very useful since it provides to the experts a class of scenarios of accidents similar to the new case treated and getting closer to the context of new case. The Case-based reasoning is then exploited as a mode of reasoning by analogy allowing to choose and to recollect one under group of historical cases that can help in the resolution of the new case introduced by the experts. Process-Oriented Case-Based Reasoning (PO-CBR) is a growing application area in which CBR is used to address problems involving process data in a variety of specialized domains. PO-CBR systems often use structured cases. Our approach is characterized by a two-phased retrieval strategy. A first phase consists in retrieving a set of cases to be considered (a class of cases most similar to a problem to resolve). In a second phase, a more fine grained strategy is then applied to the pool of candidate cases already selected by the mean of similarity measures. This approach can enhance the process of retrieving cases compared to an exhaustive case-by-case comparison.**

Keywords-componentSecurity of transport; Artificial Intelligence; Ontology; Case-Based Reasoning; Resolution scenario; Scenario of accident.

I. INTRODUCTION

A **scenario of accident** or of incident is seen as a dynamic evolution of a habitual situation in the transportation system until the point of attaining a risky situation. This undesirable situation in terms of security is then exploited to reinforce the security of the system by acting on functions of security and adequate automatisms. The experts of the certification have to study the transportation system in order to bring out its compliance with recognized guideline of railway security and have to decide to accord him an interim or final amenity before its bet in site.

In another side, the constructors of the transportation system put in contribution the report of security study given by the experts in order to improve the security level of their transportation systems. The simulation of accidents is an essential activity of the experts that puts in contribution the mind of analysis and of synthesis of the human experts. An analysis of the different components of the transportation system and their modes of faults allows envisaging the summed up faults of the system. These summed up faults when injected into the model of Petri allow to the experts to make scenarios of accidents. All scenarios made by the experts allow having an idea of the compliance degree of the transportation system offered by the constructor in norms and guidelines of security. Currently a model of feasibility of the approach based on case is under way. This model uses ontology of the domain of the transport allowing the unification of the vocabulary used in analysis of railway security.

II. CONCEPT OF ACCIDENT SCENARIO

A scenario of accident represents an unsecured situation concerning automated transport system. This potential insecurity should be solved by proposing a solution to be adopted in order to overcome the risk that it represents (example: collision between two underground oars or derailment of the way etc.). A scenario is a combination of circumstances that can lead to a danger. It is jointly described by attributes of situation that give an idea of insecurity and attributes of solution to adopt to ruin or to reduce the risk of insecurity. These descriptors take a census of several parameters, which report a risk / functions of security / actors / the geographical zone / breakdowns touching a part of the system / solutions adopted to ruin risk (See Figure 1.)

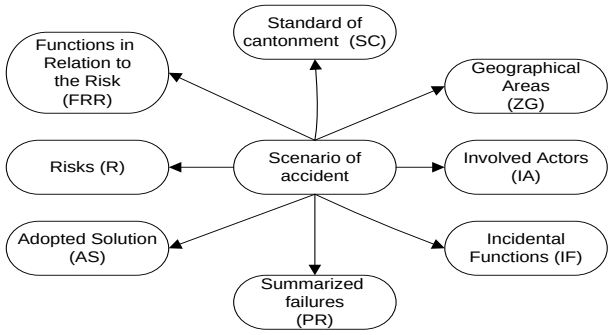


Figure 1. The parameters of static description of an accident scenario

III. PREVIOUS WORKS

This problem of analyzing security was already treated according to a first approach during the thesis jobs of [17] and [10]. The accomplished jobs had led, first, to the implementation of a feasibility model and to the achievement of an acquisition and classification system of the scenarios of accidents called ACASYA. In order to help the experts in their activity of valuation of accidents scenarios, ACASYA articulates around three supplementary stages (See figure 2):

- First, **classify** the scenario of accident in a predefined class with the assistance of classification algorithm. Every class of scenarios is characterized by a description composed by the most pertinent scenarios descriptors of the class.
- Then, **evaluate** the scenario offered by the constructor in reference to the membership class found in the previous stage to restrain the exploration space. The evaluation consists of testing the completeness and the coherence of a scenario of accident.
- **Generate** new unsecured situations by putting in contribution the superfluous and/or missing elements of description discerned in valuation by EVALSCA.

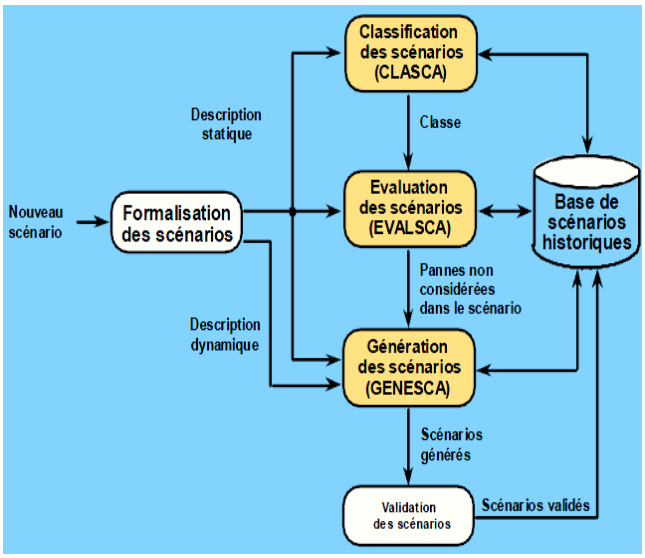


Figure 2. ACASYA System Architecture [17], [10].

The first results of the valuation of our research works applied to the domain of the analysis of security [17], allowed us to try so to identify a more generic model of acquisition of knowledge [20], and a stage step more general for the resolution of problem [19].

IV. OUR CURRENT RESEARCH FOCUS

In the previous works, several borders were disclosed and deserve one particular attention for our current works:

- ACASYA limited itself to exploit the static description of the scenario of accident. Although, this description is related easily to an example of training (chip of pair < Attribute/value>), it does not allow only to report the wealth of an accident which makes intercede several material/software and human actors and implicates several automatism and functions of security, etc. **It would be then natural to take advantage from all the forms of description of scenario: dynamic, textual and graphic description**.
- The static aspect of the scenarios of accidents represents the general context of the problem put down by accident as well as precautionary or corrective measures (adopted solutions). However, it gives no idea of the holding of scenarios in time and in space. It is only the dynamic description, which takes responsibility for it. **This dynamic description takes a well part in simulation**.
- The language of description used in the static panel of the scenarios of accidents misses a total coherence and do not consults any terminology or vocabulary well defined in the domain. It is then opportune to supplement this language in order to **set up a domain ontology**, which would raise all used concepts, relations between concepts and authorities of these concepts in the domain. This ontology would be a reference frame to describe the scenarios of accidents and to measure their completeness and pertinence of syntactic and semantic angle.
- ACASYA operates in classifying theoffered scenarios of accidents by the constructor to report its acceptability to a group of classes of scenario predefinedby the experts. ACASYA leads by training for each from the predefined classes **to a characteristic description of the class** under the form <Attribute/Values/frequency of appearance in its class>. One brings closer then according to a similarity measure to the description of a class of scenarios to a new scenario (static description): the one for whom it represents the best similarity score according to the common attributes/values. It would be more opportune to go through the different scenarios of the class in order to measure similarity inter-scenarios. This, to spot the best historical scenarios of accident (the most similar to the new scenario) stocked in the database. **This include a case-based reasoning (CBR) and not a purely inductive mechanism. Indeed, we consider CBR to be the means of the return of experience (Rex) in the field of the analysis of security.**
- Another problem in the existing system is because of knowledge representation model (scenario) and the problem resolution model (induction): these are two separated elements and remind a classic approach of treatment. **Moreover, in an approach based on knowledge, it turns more and more in research to include the method of resolution of problem.**

The following sections are then going to be dedicated to introduce the model of case as we envisage it as well as at jobs undertaken in the group to set up the ontology of domain. This model of case is envisaged to be honest with dynamic description and favors the simulation of case of accidents to report notably the analysis of security. The final section will be

reserved for the presentation of the general steps based on case adopted for the strengthening of the analysis of security and to help in the decision of the experts.

V. CASE-BASED REASONING (CBR)

The Case Based Reasoning is a type of reasoning in Artificial Intelligence AI. Case based reasoning means remembering past situations similar to the current situation and by these situations to help resolve the current situation. The case based reasoning (CBR) is a form of reasoning by analogy [10, 20]. The analogy searches for cause and effect relation in past situations and transfer them to the current situation. The case based reasoning research only looks for similarities or proximity relations between past situations and the current situation. The C.B.R. Considers reasoning as a process of remembering a small set of practical situations: the cases, it bases its decisions on the comparison of the new situation (target cases) with the old (reference cases). The general principle of CBR is to treat a new problem (target case) by remembering similar past experiences (base case) [10, 20]. This type of reasoning rests on the assumption that if a past experience and new circumstances are sufficiently similar, then everything can be explained or applied to past experience (base case) and remains valid when applied to the new situation which represents the new problem to solve. From a very global view, the CBR uses a basis of experience or case, a mechanism for searching and retrieving similar cases and an adaptation mechanism and evaluation solutions of selected cases emanating in order to solve the specified problem [28] (See Figure. 3).

A "case" is a structured representation of a composed history crossed or imagined [15]. According to [24], a case can be defined as a group of contextual knowledge teaching a lesson. A CBR process passes by the four following stages:

- **Recall stage**: further to the development of a target case, a research is performed in the database of cases to find those likely to resolve the problem.

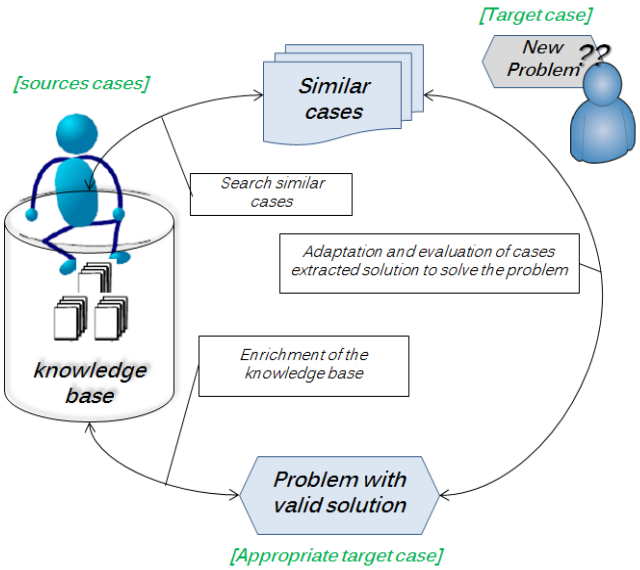


Figure 3. Cycle case based reasoning (CBR) [18].

- **Adaptation stage**: This stage is necessary because the recollected case is never strictly the same as the new case. It consists in changing the solution of the recollected case to consider difference between specifications of problems. It asks for specific knowledge of adaptation that is always easy neither to model nor to implement.
- **Revision stage**: After its generation by the system, offered solution is tested in order to be validated during the revision stage, which is generally external to the system.
- **Memorization or training stage**: This stage of the CBR cycle tries to improve knowledge at the origin of chess met by solution and to enrich the database of case with new resolute cases.

Three CBR representation of case exist in the art state.

A. The structural model

In this model, all important characteristics that describe a case are beforehand determined by the system designer. The similarity between two cases is measured according to the distance between the values of the same attributes. This distance is often estimated by Euclidean measures or by Hamming measures. The total similarity between two cases is habitually estimated by a sum balanced of the similarity of each attributes. All works on the adaptation of case are led as part of the structural model.

B. The conversational model

In the structural model, a problem must be completely described before the research in the database of case starts. Moreover, this requirement presupposes an expertise of the application area, which is not the case with all the users of CBR systems. As its name points, the interactive model bets on correlation between the user and the system (from which the notion of "dialogue") to define problem progressively, to be solved and to choose the most appropriate solutions. In the interactive diagram, the correlation between the system and the user is made as follows:

- The user gives to the system a short textual description of the problem that will be solved and the system calculates the similarity between this description and the "problem" section of cases. The system offers then to the user a series of questions.
- The user chooses the questions that he wants to answer. For every answer given by the user, the system reassesses the similarity of each of cases. The questions that did not have received an answer are introduced in descending order of priority.

C. The textual modal

The practitioners of the domain have determined the borders of the structural approach and so they offered other models to spread its application in various domains. In this approach, textual cases are "not-structured" or "semi-structured". They are "not-structured" if their description is completely "Free-text". They are "semi-structured" when the text is cut up in portions labeled by descriptors such as

"problem", "solution", etc. Textual CBR differs from that structural in whom texts are simply character strings.

VI. CONCEPT OF ACCIDENT CASE

Certification Experts use the terminology "scenario" to express an accident. This concept is richer than the concept of cases found in literature. Indeed, an accident brings all the three complementary descriptive views of the accident (See Figures 5,6):

A. Textual descrption versus textual model of cases

It is about a text, which explains the holding of accident. For example, we give the description given by the experts for the scenario N°34 in the database (See figure 4).

DESCRIPTION DE *L'ÉCHEC D'EFFACEMENT D'UN ÉLÉMENT SECOURU APRÈS DÉS INITIALISATION*

– Élément A est devenu muet (il ne dialogue plus avec le pilote automatique ou PA) et il est pris en charge par le PA du tronçon n nommé PAn qui est en cours de lancer une initialisation par parcours de l'élément B en *conduite manuelle* (CM).
– L'élément B accoste l'élément A et se met en *conduite manuelle secourue* (CMS).
– Alarme de *désinitialisation*.
– Solution : il faut vider la section de tout élément avant de procéder à une initialisation.

Figure 4. Textual description of an accident scenario.

This description is similar to the textual model of case in which, textual cases are "not-structured". Their description is completely in "Free-text". However, it is possible in this text to spot the problem part in sentences 1) 2) 3) and the solution part in 4).

B. Static description versus strustural model of cases

It is a group of descriptive parameters of a scenario in the form of a chip <Attribute/Values>. Then at this level, it takes a census of several characteristic parameters, which give an account of the accident description. Static description is similar to the structural model of case that was well and strongly used in the ACASYA system.

C. Dynamic description versus conversational model of cases

It is a dynamic view of the sequential holding of an accident in the time and in the space. This representation calls Petri in as a formalism in which places correspond to states, transitions correspond to the possibilities of evolution from a state to another. This description is related to the interactive model of case. In the structural model, the user must have in priori a good idea of all factors that can influence the resolution of his problem. However, for some problems, it is difficult to determine beforehand the aspects of situation. As its name points it out, the interactive model bets on correlation between the user and the system to define the problem progressively and to choose the most appropriate solutions. Dynamic description and its execution from a given situation (tokens put in some places) could be made only thanks to a dialogue between the system and the expert in order to bring some change in the system to reach an accident. The purpose of the expert is to generate scenarios likely to improve the exhaustiveness of the security file of the railway transportation system.

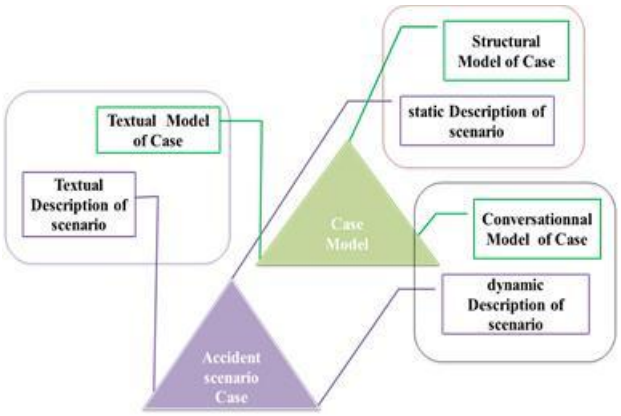


Figure 5. Case model of accident scenario.

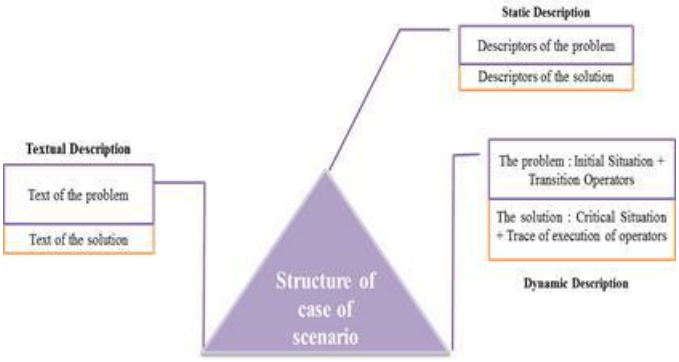


Figure 6. Structure of accident scenario case.

VII. DEVELOPING A DOMAIN ONTOLOGY

A. Presentation

Reference [4] proposes the following definition: "an explicit specification of conceptualization" that is so far the definition most cited in the literature in artificial intelligence. This definition was modified slightly [26] as "formal specification of a shared conceptualization".

These two definitions are resumed in [27] as "a formal and explicit specification of a shared conceptualization".

- Formal: the ontology should be machine readable, which excludes natural language.
- Explicit: the explicit definition of the concepts and constraints of their use.
- Conceptualization : the abstract model of a real world phenomenon by identifying the key concepts of this phenomenon.
- Shared: the ontology is not the property of an individual, but it represents a consensus accepted by a community of users.

B. The components of ontology

It is possible and even advisable to use the plural to refer to the notion of ontology to reflect the many facets that it covers. According to [27] there are several types of anthologies according to the model domain and possibly the tasks for which they are designed.

Ontology can be seen as a lattice of concepts and of relations between these concepts intended to represent the objects of the world under a comprehensible form at the same time by the men and by machineries. Ontology consists of concepts and relations as well as properties and axioms [2]. We drew inspiration from works of [2] to offer the model of case. A case in our system [5] represents a rail accident (See figure 7).

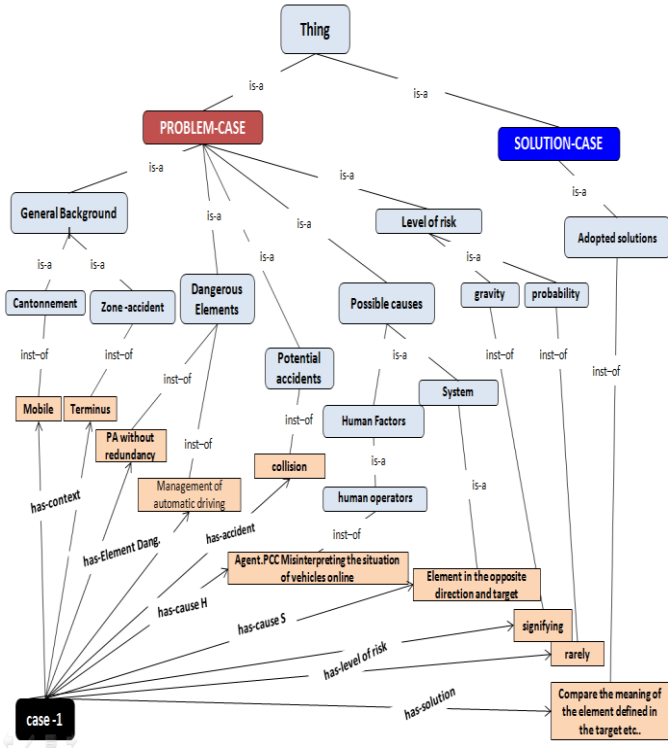


Figure 7. Case model in the domain ontology.

VIII. A UNIFIED MODEL FOR REPRESENTATION AND PROBLEM SOLVING TO KNOWLEDGE REUSE

This model allows homogenization between the knowledge representation and its exploitation. It is represented by a "scenario of resolution" of problem, which is defined as a process of problem resolution. This one is activated by an initial problem to be solved. Moreover, this model is described, since the beginning of problem up to its completion by stressing the different instants of resolution or else to the various knowledge and no mobilized views [19].

A notion of "scenario of resolution" is described according to two levels of abstractions in downward direction (See figure 8). Every level illustrates a progress in the process of resolution of the problem. "Scenario of resolution" is visible by its resolution elements, which is a part of the object view of the resolution because this element gives the trace of real resolution; on the contrary, this one would be deficient to illustrate wealth in knowledge of a scenario.

Indeed, it is the contribution of all the units of resolution, which get involved in process and give information about the totality of the scenario (from start to finish).

A. *The model of resolution*

This model (See Figure 8) theoretically shows the bases used in the process of problem solving. This is a view oriented "class of sub-problems" and gives methodical materials, procedures and knowledge necessitated to achieve good resolution. The view model is based on three axes:

- The first line: "class problem" informs on the family of problems concerned with the resolution. Indeed, knowledge modeled at this level is not tied to a specific problem, but to a class of similar problems. This is the Framework of problem solving.
- The second line: "Knowledge resolution" provides a formal description of knowledge that can be used in a possible resolution. This knowledge is both static (attributes, values) and dynamic (functions, methods of resolution, etc.).. Methods involved in a dynamic description can have an immediate character (offer an immediate solution), or a reducing character. A method has a reducing character when it divides a given problem into several sub-problems most basic.
- The third axis "Viewpoint" reflects a mode of observation of the problem. This mode allows characterizing the space of attributes and their rating scale. If knowledge resolution represents a formal theory for a class of problems, the viewpoint reflects sensitivity to a particular method given. This is an action plan to extract knowledge from formal resolution.

B. *The object of resolution*

It shows another view of the complementary model to previous resolution. Indeed, this view is just another instantiation of "model resolution" as a particular problem. This view is also three-dimensional (See Figure 8). It is based on three pillars which are the image of the three pillars of the view model:

- Pillar "Problem" which provides information about a problem or issue in particular to analyze and decompose or solve immediately;
- Pillar "Background Resolution" :Specifically, the context is represented by the set of logical and / or physical constraints imposed in problem solving. Therefore, constraints represent a "bias" in the resolution process. A constraint can be expressed by one or more logical conditions of the form <attribute value comparison>. It is important to note that the context of resolution significantly affects the view. Indeed, it is built around the operating point evidenced by the actual context of the resolution. In the specific problem of transport systems security analysis, the context of resolution is represented by "geographical zone of accident", "Human and material actors involved", "causes of accident",…etc.
- The last pillar, called "resolution" informs on the effective resolution of the problem. This is the instantiation of knowledge by solving real parameters under particular problem. This instantiation results

inthe production of a solution. This solution is characterized by the values and attributes set obtained after leveraging of useful methods for solving the class (i.e.: methods specified in the view).

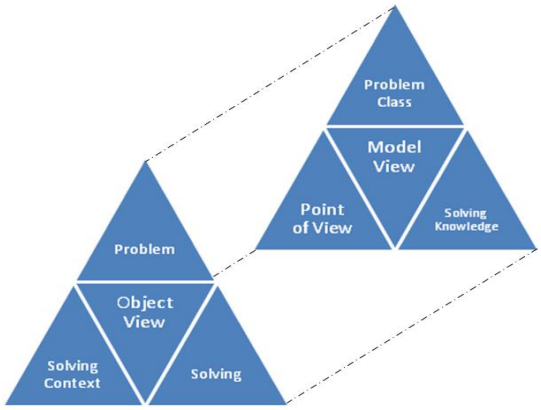


Figure 8. Complete diagram of a problem solving scenario

IX. CBR APPROACH FOR THE EVALUATION OF CASES OF ACCIDENTS

We can note that a scenario is richer than a case as it regroups three models of case: The structural, the conversational and the textual model. In literature, it is often a matter to adopt only one of these three models. According to the multiform aspect of a scenario, it is judicious to exploit these three models and then to improve the analysis of security and not to limit itself to the static description strongly used in ACASYA to the detriment of others. We illustrate our offered approach as in figure 8.

A. Knowledge formalization

It is an upstream stage of our approach. It is necessary to construct a platform of data for the analysis of problems. It consists in a collection of information and of domain knowledge across interviews with the experts of domain, documentary resources, and relational databases. Knowledge instituted during this stage are formalized under a database containing textual, structural and conversational cases, and a domain ontology which structures and formalizes knowledge exit from this stage.

B. Classifiying the accident cases

The classification in ACASYA showed a feedback of exploitable information. This classification segment all cases into contextual pieces sharing descriptions of problems. These contextual pieces represent a general description of a class of problem under the formalism <Attribute/value/frequency of appearance in a class>. In this sense, each class of accident scenarios is represented by both:

- **Extensive description** which regroups all scenarios belonging to the class.
- **Intensive description** which is materialized by a characteristic description. This one regroups all the descriptors which are more representative in the class: descriptors frequently expressed in the scenarios of this class.

This increases the deducted knowledge during the previous stage and supervises the analysis space in a simple class.

C. Recall of accident case

This stage is very important to focus in the more important and relevant cases instead of processing all cases included in base of cases [14], [23].

In the offered approach, we adopt a Case-Based Reasoning. This reasoning is applied in the membership class of the problem case. This one is done in order to restrain the treatment of the process in a simple class of cases. The first Stage of this reasoning is the recall of past accident cases.

It is about a research stage, which goes through all the cases of accident stocked in the class, which was spotted by ACASYA in order to identify the closest case with new constructive case in reference. A calculation of the total similarity between cases is made because of the notion of local similarity (between attributes) and by referring to the domain ontology, which is a form of normalization of all attributes in domain.

Similarity measure is calculated according to following functions:

$$Sim_{global}(Case_{target},Case_{source})=\sum^{i=n} Sim_{local}(A_{target,i},A_{source,i})*W_i$$

With:

Wi= The weight of Attribute Asource,i and Atarget,i

Casetarget= represent the case offered by the constructor.

Casesource= any case of the class identified by ACASYA.

n: Number of attributes of the part problem and not solution.

$$Sim_{local}(A_{target,i}, A_{source,i})=\sum^{k=p} \partial ik$$

With:

$\partial ik = 1$ if the value of Vk de Atraget,i=Vk de Asource,i

$\partial ik = 0$ if the value of Vk de Atarget,i=Vk de Asource,i

P = Number of descriptive values of attribute.

D. Adaptation of accident case

The evaluation of an accident scenario is to test the completeness and coherence. Completeness informed about the fact that the accident evaluated includes all descriptors needed for a good definition. Consistent informs on the integrity of the accident scenario evaluated and the fact that it has no unnecessary or redundant descriptors. The assessment leads to the detection of causes of the accident and the proposal of solutions to overcome this insecurity. This stage could be done according to three phases of treatment:

- Partitioning of descriptors
- Generation of adaptation rules
- Adaptation of the most similar case

1) Partitioning descriptors

The extension of the evaluation method to the set of descriptors has led to a partitioning them in 3 families (fig. 9):

TABLE I. FAMILIES OF DESCRIPTORS IN ACCIDENT SCENARIO.

Families of descriptors	List of Attributes	List of Possible Values	Selected Values
Symptoms Descriptors of accident Scenario	**Principle of Mobile**	Fixed Canton	x
		Mobile Canton	
	Risk	Collision	x
		Derailment	
		Others	
	Functions of Security	Automatic Driven Management	x
		Docking of trains	
		Others	
	Geographical Zone of Accident	Terminated station	
		Line	x
		Others	
Causes Descriptors of Accident Scenario	**Involved Actors**	Number of trains	
		Operator at PCC	
		Others	
	Secondary Functions of security	Management of Itineraries	
		Instructions	x
		Others	
	Summarized Failures	PR6 : Impossible Start-up of train	x
		PR2: Train with collision Transmitter Failure	x
		PR17: Alarm masking by docking	x
Cure Descriptors of Accident Scenario	**Adopted Solutions**	SA12 : Check effective docking of Train	x
		Others	

- Descriptors like "**symptoms**" are used to characterize the problem and the context in which the story unfolds insecurity. They concern such descriptors:

(Risk = Collision) (Principle Mobile = township fixed) (Location = Terminus) (Actors involved = mobile operator) ... etc...

- The type descriptors "cause" that define human failures and / or material leading to danger. Failures elementary incorporated more generic groups designated "Summarized Failures" (PR) such as:

(Summarized Troubleshooting = invisible element of the driving area Full Auto).

- The type descriptors "cure" solutions that are intended to maintain or restore the operating system and therefore limit the impact of failures. In the context of transport, these remedies are called Adopted Solutions (SA). The remedies are usually offered by the manufacturer of the transport system in terms of palliative risk such as:

"Solutions Adopted" = "check list L4 is free."

2) Generation of adaptation rules

The generation of adaptation rules is made by comparing the cases of accident belonging to the same class. To do this, a learning algorithm of rules was applied. We chose the algorithm CHARADE [13] to induce a set of rules for adaptation from the class of accident scenarios identified in the previous step by the classifier. Interest CHARADE system is it does not produce a rule base isolated that is to say the rules of classification of type:

IF conjunction of descriptors THEN Class.

But it can produce outright by inductive inference, a system of rules of completion type:

IF conjunction1 of descriptors THEN Conjunction2 of descriptors.

The empirical induction system CHARADE can detect similarities between scenarios insecurity of the same class, considered as training examples, thus forming adaptation rules. The rules induced then contribute to the CBR process. The evaluation provides an overall assessment of the completeness and consistency of the scenario studied.

The analysis of the natural process of reasoning experts reveals that in fact, experts are alarmed by the symptoms that they are looking for causes and then determine remedies. Two stages may be distinguished in the expert reasoning:

- First step: symptoms suggest causes;
- Second stage: symptoms and causes require remedies;

However, some symptoms may bring in experts to consider other causes and some allow evoking others. This leads as shown in Figure 10, to enrich the previous line of reasoning.

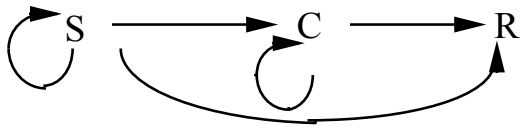


Figure 9. Natural reasoning of experts.

This analysis motivates the construction of two systems of rules:

- One Recognition Causes System (SRC) which contains rules of the form:

IF so (s) symptom (s)

THEN it (s) generally coexist (s) with such (s) other (s) symptom (s)

AND is (are) emanated from such (s) cause (s).

Example:

IF SYMPTOMS: (Risk = Collision)

AND (Security Functions = Alarm Management)

AND (Involved Actors = Itinerant Operator)

THEN SYMPTOMS: (Security Functions = Docking)

AND CAUSES: (Summarized Failures = PR17).

- One Recognition Remedies System (SRR), which includes rules of the form:

IF the presence of such (s) symptom (s)

AND IF it (s) emanate (s) from so (s) cause (s)

THEN it(s) generally coexist (s) with such (s) other symptom (s)

AND so (s) cause (s)

AND require (s) such (s) remedy (ies).

Example:

IF SYMPTOMS: (Risk = Collision)

AND (Security Functions = Alarm Management)

AND (Involved Actors = Itinerant Operator)

AND CAUSES: (Summarized Failures = PR17)

THEN SYMPTOMS: (Security Functions = Initialization)

AND CAUSES: (Summarized Failures = PR2)

AND REMEDIES: (Adopted Solutions = SA12).

The use of CHARADE system [13] has generated two induction systems with 80 rules about rules for SRC and 50 rules for the SRR. This large number of rules can be explained by the existence of rules of completion that is to say the rules go descriptors symptoms and conclude on other symptoms to complete the description.

3) Adaptation of Case

To illustrate the method of adaptation of case, let us call the Case To Evaluate as (CTE) and the Most Similar Case obtained by CBR mechanism of recall as (MSC).

According to this we obtain two cases:

- The case to evaluate named CTE is composed of three components: Symptoms (S_{CTE}); Causes (C_{CTE}); Remedies (R_{CTE}).

$CTE = S_{CTE} \cup C_{CTE} \cup R_{CTE}$ (***Case To Evaluate***)

- The Most similar case named MSC is composed of three components: Symptoms (S_{MSC}); Causes (C_{MSC}); Remedies (R_{MSC})

$MSC = S_{MSC} \cup C_{MSC} \cup R_{MSC}$ (***Most Similar Case***)

The first step is to determine relevant descriptors which are defined as descriptors envisaged both in case to evaluate and in the most similar case selected. Relevant descriptors are common descriptors listed in the case to evaluate and the historical case remembered by the CBR system. In this way, we obtain three kinds of relevant descriptors as follows:

- Relevant Symptoms labeled (S_{Rel})

$(S_{Rel}) = S_{CTE} \cap S_{MSC}$

- Relevant Causes labeled (C_{Rel})

$(C_{Rel}) = C_{CTE} \cap C_{MSC}$

- Relevant Remedies labeled (R_{Rel})

$(R_{Rel}) = R_{CTE} \cap R_{MSC}$

Three techniques are here used to adapt the most similar case selected by CBR process.

a) Deductive Adaptation

This technique uses deductive inference to obtain an adapted case in two inferences according to adaptation rules systems labeled SRC and SRR and driven by Relevant Symptoms (S_{Rel}) as follows:

Inference 1:
(S_{Rel}) *Deduction By SRC* S_{Comp1}, C_{Comp1}

Inference 2 uses descriptors obtained by inference 1:
$\{S_{Rel} \cup S_{Comp1} \cup C_{Comp1}\}$ *Deduction By SRR* $S_{Comp2}, C_{Comp2}, R_{Comp1}$

Case obtained after adaptation is:
$\{(S_{Rel} \cup S_{Comp1} \cup S_{Comp2}) ; (C_{Comp1} \cup C_{Comp2}); (R_{Comp1})\}$

The first packet of Compatible descriptors obtained by inference is labeled Comp1; the second one is labeled Comp2.

b) Abdicative Adaptation

This technique uses abdicative inference to obtain an adapted case in two inferences according to adaptation rules systems labeled SRR and SRC and driven by Relevant Remedies (R_{Rel}) as follows:

Inference 1:
(R_{Rel}) *Abdication By SRR* S_{Comp1}, C_{Comp1}

Inference 2 uses descriptors obtained by inference 1:
$\{S_{Comp1} \cup C_{Comp1}\}$ *Abdication By SRC* S_{Comp2}

Case obtained after adaptation is:
$\{(S_{Comp1} \cup S_{Comp2}) ; (C_{Comp1}); (R_{Rel})\}$

c) Abdicative and Deductive Adaptation

This technique uses both abdicative and deductive inference to obtain an adapted case in two inferences according to adaptation rules systems labeled SRC and SRR and driven by Relevant Causes (C_{Rel}) as follows:

Inference 1: (Abdication)
(C_{Rel}) *Abdication By SRC* S_{Comp1}

Inference 2: (Deduction)
$\{S_{Comp1} \cup C_{Rel}\}$ *Deduction By SRR* $S_{Comp2}, C_{Comp2}, R_{Comp1}$

Case obtained after adaptation is:

$$\{(S_{Comp1} \cup S_{Comp2}) ; (C_{Rel} \cup C_{Comp1} \cup C_{Comp2}); (R_{Comp1})\}$$

E. Revision of accident cases

After adaptation, the solution of the new structural target case is got. This got case should be reviewed by the experts aiming the validation and though to allow toping it up at the root of case and to enrich it by indexing them and by training.

F. Training

The solution got during the previous stage and validated by the expert of domain is topped up at the root of case and then to enrich it, and consequently to enrich its indexes. In addition, this resolution trace is archived in the form of a "scenario of resolution" of problem, which could be reused completely in the future.

G. Database of resolution scenario

The main contribution of this approach is the harmonization between knowledge representation and its exploitation. Moreover, this latter by using a generic model of representation and resolution offered by [19]. The projection of this model in our context was as follows:

For the object view incorporated of three pillars (Problem, Context of resolution and resolution), we divide knowledge as follows:

- Problem: it contains the symptoms description of the accident case by considering its membership class.
- Resolution context: This contained the knowledge of the context of the problem, such as Causes descriptors.
- Solving: This pillar is by equivalence of its definition contains Remedies descriptors such as the adopted solution deducted from CBR.

For the model view incorporated of three pillars (Class of problem, Point of view and knowledge of resolution), we divide knowledge as follows:

- Class of problem: it is the description of the membership class of the problem generated by ACASYA. This class is represented by extensive description in term of accident scenarios enumerated.
- Point of view: it is the modality of the problem observation, it is represented by the point of view of resolution (Structural/ or conversational/ or textual). In our first study we focus on structural CBR.
- Knowledge of resolution: it is capitalized by grouping the adaptation rules of the class (SRC and SRR), characteristic descriptions of the classes of problems, the domain ontology and the index of the accident cases used during the CBR process, etc.

H. Implementation of the offred approach

We attract in this section to introduce the model system that assess the offered approach. This model although it is still under the development stage, we can describe its general architecture and introduce some windows of the application (See Figure 11).

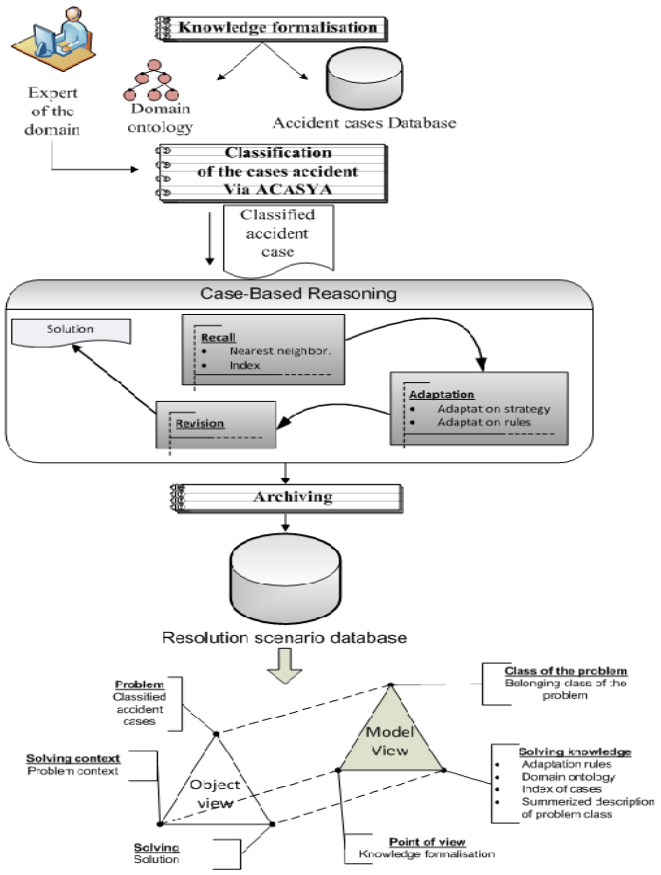


Figure 10. Complete diagram of the proposed approach.

The architecture of the ongoing model spreads out mainly on five modules:

- A preconception module, which includes the part of knowledge formalization having the goal to structure the knowledge by defining the domain ontology under an XML format, keyboarding of the domain dictionary. Expert knowledge is defined throughout by the specification of similarity parameters.
- A module of accident cases management, which allows addition, modification and removal of accident cases according to their representation models: textual, structural and conversational.
- A classification module of accident cases,same as ACASYA.
- A module represented by an engine of CBR reflecting its different stages (Recall, adaptation, revision and training).
- A module for knowledge archiving in a database of "scenario of resolution", which keeps the trace of the problem resolution as well as the used knowledge for the resolution. This scenario could be reused in the subsequent sessions of exploitation of the help system.

Finally, we introduce into this paper an example of an execution window of the system, which is under developing:

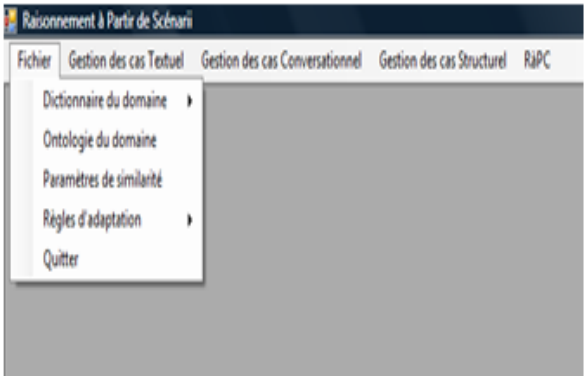


Figure 11. The main window of the system.

This window, which represents the main window of application, allows going through all the functionality described in the architecture of the system.

X. CONCLUSION

We introduced in this article the first works made concrete by our approach and notably the general architecture of the CBR system. This architecture articulates around two processes: an off-line process and other one on-line, which use different types of knowledge for the problem solving. In result, the introduced work as a part of this article follows from the problems of knowledge management of a critical domain: the security and especially, the railway accident. Although the application field constitutes in itself originality, we can stress the still generic character and the opening of knowledge model.

This article presents just a work in progress. We must emphasis our study on others methods of CBR such as textual and conversational CBR. Currently, we work on the acquisition of knowledge relating to reasoning and to the improvement of the model worked out of knowledge, especially ontology, to enrich it by concepts relating to the dynamic description of the accident scenario. It is judicious to study the recent works in the field of CBR which used ontology in different ways: to describe and to structure cases [15], to offer independent models of domain [9], CCBRONTO [6]; to calculate semantic similarity [27], to treat the heterogeneity of cases [2], etc.

REFERENCES

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first . . .”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors' names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

[1] A. Aamodt, ″A Knowledge-Intensive, Integrated Approach to Problem Solving and Sustained Learning,″. PhD Thesis, University of Trondheim, Norway (1991).

[2] A. Abou Assali, ″Acquisition des connaissances d″adaptation et Traitement de l″hétérogénéité dans un système de ràpc basé sur une Ontologie. Application au diagnostic de la défaillance de détecteurs de gaz,″. Thèse De Doctorat (2010).

[3] A. Abou Assali, D. Lenne, and B. Debray, “Ontology Development for Industrial Risk Analysis,”. In IEEE International Conference On Information, Communication Technologies : From Theory ToApplications (ICTTA’08), Damascus, Syria, April 2008c.K. Elissa, “Title of paper if known,” unpublished.

[4] A. Maalel, andH.Hadj-Mabrouk,″Contributionof CaseBasedReasoning(CBR) in the Exploitation of Return of Experience. Application to Accident Scenarii in Rail Transport,″. SIIE'2010, 3rd International Conference OnInformationSystems AndEconomic Intelligence,18-20february2010,Sousse,Tunisia.

[5]] A. Maalel, H. Hadj Mabrouk,L. Mejri and H. Ben Ghezela, “Development of an Ontology to Assist the Modeling of an Accident Scenario: Application on Railroad Transport,”. Journal Of Computing, Volume 3, Issue 7, July 2011.

[6] B. Dìaz-Agudo,and P.A. Gonzàlez-Caler, "CBROnto : a task/method ontology for CBR,”. Procs. Of The 15th International FLAIRS, 2 :101–106. 2002.

[7]]C. Roche,″ Le terme et le conceptfondements d’une ontoterminologie,″. PremièreConférence Terminologie & Ontologie : Théories et Applications. TOTh 2007.

[8] C. Roche, ″Ontology: a Survey,″. The 8th Symposium On Automated Systems, IFAC. Goteborg, Sweden, September 22-24, 2003.

[9] G. Gauchía, B. Díaz-Agudo, and P.A. González-Calero,“Ontology-Driven Development of Conversational CBR Systems,”. In Advances In Case-Based Reasoning, Proceedings Of The 8th European Conference (ECCBR’06), volume 4106, pages 309–324, Turkey. Springer Berlin / Heidelberg. ISBN 978-3-540-36843-4, (2006).

[10] H. Hadj-Mabrouk, ″Acquisition et évaluation des connaissances de sécurité des systèmes industriels. Application au domaine de la certification des systèmes de transport guidés,”. Thèse d″Habilitation à Diriger des Recherches, Université De Technologie De Compiègne 1998.

[11] J. Charlet,.″ Constructionderessources terminologiquesou ontologiquesà partirde textes,″, Cadre Unificateur, (2004).

[12] J.A. Recio-Garcìa, B. Dìaz-Agudo, P. González-Calero, and A. Sanchez, ″Ontology based CBR with COLIBRI,″. Applications And Innovations In Intelligent Systems, 14 : 149–162, 2006.

[13] J.G. Ganascia,″L'apprentissage symbolique,″, Encyclopédie De La Communication - PUF, S/direction de Lucien SFEZ, 1991.

[14] J. Kendall_Morwick, D. Leake, “Facilitating representation and retrieval of structured cases : Principles and toolkit”, Information Systems (2012) in press.

[15] J.Kolodner, ″Case-Based Reasoning,″.*Morgan-Kaufmann Publishers*, Inc., 668 pages, 1993.

[16] J. Lieber, M. d’aquin, S. Brachais, and A. Napoli, ″Une étude comparative de quelques travaux sur l’acquisition de connaissances d’adaptation pour le raisonnement à partir de cas″. In Actes du *12*ème atelier de raisonnement à partir de cas (RàPC’*04*), 2004.

[17] L. Mejri,″Une démarche basée sur l″apprentissage automatique pour l″aide à l″évaluation et à la génération de scenarios d″accidents. Application à l″analyse de sécurité des systèmes de transport automatisés,″. Université de valenciennes, 6 décembre 1995.

[18] L. Mejri., A. Maalel. and H. Ben Ghezala, ″Simulation de cas d’accidents pour l’aide à la décision.Application à la sécurité des systèmes de transport ferroviaires automatisés,″. Atelier Aide EGC 2012, 12ème Conférence Francophone Sur l’Extraction Et La Gestion Des Connaissances, 31 janvier au 3 février 2012, Bordeaux.

[19] L. Mejri, H. Hadj mabrouk, and P. Caulier,″Un modèle générique unifié de représentation et de résolution de problème pour la réutilisation de connaissances. Application à l'analyse de sécurité des systèmes de transport automatisés,″. Revue Recherche, Transport Sécurité, RTS N°103, 2009.

[20] L. Mejri, and P. Caulier, ″Formalisation of a scenario concept to dynamic problem solving,″, 24th European Annual Conference on Human Decision Making and Manual Control EAM, Athens, 17-19 October 2005.

[21] M. H. Mazouni, ″Pour une Meilleure Approche du Managementdes Risques : De la Modélisation Ontologique duProcessus Accidentel au Système Interactif d'Aide à laDécision″ ,Thèse de doctorat 2008.

[22] M. Lopez, A. Gomez-Perez, and M. Rojas-Amaya, ″Ontology's crossed life cycle″. The 12th International Conference On Knowledge Engineering And Knowledge Management, EKAW-2000.

[23] M. Minor, S. Montani, "Preface", In Proceedings Of The ICCBR, Workshop on Process Oriented Case Based Reasoning, 2012.

[24] M. Ushold,and M. Gruninger, ″Ontologies: principles, methods, and applications″. Knowledge Engineering Review, 11(2):93–155, 1996.

[25] N. Friedman, and D. McGuiness, ″OntologyDevelopment″, 101: A Guide to Creating Your FirstOntology. 2001.

[26] T. Gruber, "A translation approach to portable ontology specifications.Knowledge".AcquisitionJournal. 5:199 -220, 1993.

[27] W.N. Borst, ″ConstructionofEngineering Ontologies″.Thèsede doctorat, Center forTelematicaandInformationTechnology,Université deTweenty, Enschede, 1997.

[28] W.S. Han, M.D. Pham, J. Lee, R. Kasperovics, and J.X. Yu, "igraph in action : performance analysis of disk based graph indexing techniques, In Proceedings Of The International Conference On Management Of Data, SIGMOD'11, New York, NY, USA, ACM (2011) 1241-1242.

AUTHORS PROFILE

MejriLassaâd
Researcher in Laboratory of Automatic,Industrial/HumanComputer Sciences.LAMIH CNRS UMR 8530.Valenciennes – France.
Dr In Computer Sciences - Faculty of Sciences In Bizerta–Computer Science Department.
University of Carthage
7021 JARZOUNA –BIZERTE – TUNISIE
Tél. : 216-98-362222 - Fax : 216-72-590566.-
E_mail : *mejrilassad@yahoo.fr*

MadiSofian
Researcher in RIADI Laboratory of ENSI (National School of Computer Sciences), Tunis – Tunisia.
Manouba University – Tunis – Tunisia
P.O. Box 88 Hached Gabes 6001 Tunisia
Tél. : 216-22-793737 - E_mail : sofian.madi@gmail.com

Ben GhezalaHenda
Directorof RIADI Laboratory - ENSI (National School of Computer Sciences). Tunis – Tunisia.
Professor in Computer Sciences -ENSI (National Shool of Computer Sciences), Computer Science Department - Tunis–.Tunisia
Manaouba University– TUNISIA.
E_mail :hhbg.hhbg@gnet.tn

Prediction Method for Time Series of Imagery Data in Eigen Space

Validity of the Proposed Prediction Metyhod for Remote Sensing Satellite Imagery Data

Kohei Arai
Graduate School of Science and Engineering
Saga University
Saga City, Japan

***Abstract*—Prediction method for time series of imagery data on eigen space is proposed. Although the conventional prediction method is defined on the real world space and time domains, the proposed method is defined on eigen space. Prediction accuracy of the proposed method is supposed to be superior to the conventional methods. Through experiments with time series of satellite imagery data, validity of the proposed method is confirmed.**

Keywords-prediction method; eigen value decomposition; eigen space; time series analysis.

I. INTRODUCTION

There are conventional prediction methods which allow prediction of future imagery data by using the acquired imagery data in the past [1]-[6]. The conventional prediction method is defined on real world space and time domains, in particular, is defined as an auto regressive model. Namely, the conventional prediction model is using relations among the time series of data which is described with stochastic processes such as Markov model[1]. Therefore, prediction accuracy is not good enough when the time series of data has an irregularity in time domain.

On the other hands, it is possible to create eigen space [6] with time series of data. Namely, time series of data can be projected onto eigen space from the real world space and time domains. Time series of data behavior can be well described on the eigen space rather than the real world space and time domains in particular for the irregularly changed data. The conventional prediction models are based on autoregressive model, or moving average model. Therefore, it is difficult to predict a future data using previously acquired time series of data for irregularly varied data. Meanwhile, it is expected that prediction accuracy of the proposed method is better than that of the conventional method.

Through experiments with GMS/VISSR images[2] which are acquired every one hour, comparative study on prediction accuracy between the proposed and the conventional methods is conducted. The experimental results show advantage of the proposed method in terms of prediction accuracy.

[1] http://en.wikipedia.org/wiki/Markov_model
[2] http://www.tric.u-tokai.ac.jp/tsic/egms.html

The following section describes the proposed prediction method for time series analysis followed by experimental results. Then conclusion is described together with some discussions.

II. PROPOSED METHOD

A. Data Description on Eigen Spece

Eigen space is formed with the bases of eigen vectors determined in conjunction with eigen values of variance-covariance matrix derived from the vectors of acquired imagery data as vectors. The vectors are projected in the eigen space, then time series of imagery data can be represented as the vector in the eigen space.

M of imagery data, $\{\mathbf{f}^i | i = 0, \cdots, M\}$, then imagery data matrix, **F** is defined as the following equation,

$$\mathbf{F} = \left[\mathbf{f}^1 - \mathbf{c}, \cdots, \mathbf{f}^M - \mathbf{c}\right] \tag{1}$$

where the time series of imagery data,

$$\mathbf{f} = [f_1, f_2, \cdots, f_N]^T$$

in concern and the averaged images,

$$\mathbf{c} = \sum_{i=1}^{M} \mathbf{f}^i .$$

Then variance-covariance matrix is defined as follows,

$$\mathbf{Q} = \mathbf{F}\mathbf{F}^T \qquad (N \times N) \tag{2}$$

and eigen values of Q is expressed as follows,

$$\lambda_i \mathbf{e}_i = \mathbf{Q}\mathbf{e}_i$$

where e_i is defined as follows,

$$\mathbf{E} = [\mathbf{e}_1, \cdots, \mathbf{e}_N]$$
$$\mathbf{e}_i = [e_{i1}, e_{i2}, \cdots, e_{iN}]^T$$

The well known contribution factor is defined as follows,

$$W_k = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{N} \lambda_i} \geq T_s$$

Then F can be expressed with the first k eigen vectors,

$$\hat{\mathbf{E}} = [\mathbf{e}_1, \cdots, \mathbf{e}_k] \tag{4}$$

By using equation (4), time series of imagery data can be projected onto eugen space based on the following equation,

$$\mathbf{\Phi}^i = \hat{\mathbf{E}}^T(\mathbf{f}^i - \mathbf{c}) \tag{5}$$

That is same thing for representation of whole time series of imagery data by using all of the eigen vectors as follows,

$$\begin{aligned} \hat{\mathbf{\Phi}}^i &= \mathbf{E}^T(\mathbf{f}^i - \mathbf{c}) \\ \hat{\mathbf{\Phi}}^i &= \left[\phi_1^i, \cdots, \phi_N^i\right] \end{aligned} \tag{6}$$

The prediction method proposed here is based on auto regressive model with the projected time series of imagery data onto eigen space.

B. Auto Regressive Model

By using the following n-1 of time series of data,

$$\{x(s) | s = 1, 2, \cdots, n-1\}$$

n-th data can be predicted as follows,

$$\hat{x}(s) = \sum_{m=1}^{M} a(m)x(s-m) \tag{7}$$

where

$$\epsilon(s) = x(s) - \hat{x}(s) \quad (s = 1, 2, \cdots, M) \tag{8}$$

is called prediction error. All the coefficients for equation (7) can be determined by minimizing the following mean square error,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{s=1}^{N} \epsilon^2(s)$$

Thus autoregressive model can be represented as follows,

$$x(s) = \sum_{m=1}^{M} a(m)x(s-m) + \epsilon(s) \quad (s = 1, 2, \cdots)$$

Stationery condition can be expressed as follows,

$$1 - a(1)z - a(2)z^2 - \cdots - a(M)z^M = 0$$

It is called characteristic equation of autoregressive model.

For the prediction error, the following three equations are well known,

$$\begin{aligned} &E\left[\epsilon(s)x(s)\right] = E\left[\epsilon^2(s)\right] = \sigma^2 \\ &E\left[x^2(s)\right] = \textstyle\sum_{m=1}^{M} a(m)E\left[x(s-m)x(s)\right] + \sigma^2 \\ &E\left[x(s)x(s-k)\right] = \textstyle\sum_{m=1}^{M} a(m)E\left[x(s-m)x(s-k)\right] \end{aligned}$$

Together with the following three autocorrelation functions,

$$\begin{aligned} &R_{xx}(0) = \textstyle\sum_{m=1}^{M} a(m)R_{xx}(m) + \sigma^2 \\ &R_{xx}(k) = \textstyle\sum_{m=1}^{M} a(m)R_{xx}(|k-m|) \quad (k = 1, 2, \cdots) \\ &R_{xx}(k) = \textstyle\sum_{m-1}^{M} a(m)R_{xx}(k-m) \quad (k \geq M) \end{aligned}$$

As of $R_{xx}(-k) = R_{xx}(k)$, then

$$\begin{aligned} &R_{xx}(0) = \textstyle\sum_{m-1}^{M} a(m)R_{xx} + \sigma^2 \\ &R_{xx}(k) = a(k)R_{xx}(0) + \textstyle\sum_{j-1}^{M}(a(k+j) + a(k-j))R_{xx}(j) \end{aligned}$$

Therefore, if $\mathfrak{a}_{\mathfrak{M}} = (\mathfrak{a}(\mathfrak{1}), \mathfrak{a}(\mathfrak{2}), \cdots, \mathfrak{a}(\mathfrak{M}))_{\mathfrak{L}}$ and σ^2 are given, then probability density function is reduced as follows,

$$f(x|\sigma^2, a_M) = h(x_M|\sigma^2, a_M) \prod_{s=M+1}^{N} g(x(s)|\sigma^2, a_M)$$

In order to determine the parameters in the probability density function,

$$\begin{aligned} &\log f(x|\sigma^2, a_M) \\ &= \log h(x_M|\sigma^2, a_M) + \textstyle\sum_{s=M+1}^{N} \log g(x(s)|\sigma^2, a_M) \\ &= \log h(x_M|\sigma^2, a_M) + (N-M)\log(2\pi\sigma^2)^{-\frac{1}{2}} \\ &\quad -\frac{1}{2\sigma^2}\textstyle\sum_{s=M+1}^{N}\{x(s) - \sum_{m=1}^{M} a(m)x(s-m)\}^2 \end{aligned} \tag{9}$$

can be used. By maximizing the logarithmic function of the probability density function, then the parameters can be determined. It, however, is time consumable. Therefore the following Yule Walker method or the following least square method is used to be used.

C. Least Square Method

If N is much greater than M, then the first term of the equation (9) is negligible. Thus, coefficients of the autoregressive model of equation (10) can be determined.

$$\sum_{s=M+1}^{N} \{x(s) - \sum_{m=1}^{M} a(m)x(s-m)\}^2 \tag{10}$$

σ^2 can be determined with maximum likelihood manner,

$$\frac{N-M}{2\sigma^2} = \frac{1}{2(\sigma^2)^2} \sum_{j=M+1}^{N} \{x(j) - \sum_{m=1}^{M} a(m)x(j-m)\}^2$$

then

$$\sigma^2 = \frac{1}{N-M} \sum_{s=M+1}^{N} \{x(s) - \sum_{m=1}^{M} a(m)x(s-m)\}^2$$

Thus

$$\log f(x|\hat{\sigma}_M^2, \hat{a}_M) = -\frac{N-M}{2}\log 2\pi\hat{\sigma}_M^2 - \frac{N-M}{2}$$

D. Yule Walker Method[3]

Assuming

$$x(0), x(-1), \cdots, x(1-M), x(n+1), \cdots, x(n+M) = 0,$$

then

$$\log f(x|\sigma^2, a_M) = N \log(2\pi\sigma^2)^{-\frac{1}{2}} - \frac{1}{2\sigma^2} \sum_{s=1}^{N+M} \{x(s) - \sum_{m=1}^{M} a(m)a(s-m)\}^2 \quad (11)$$

The Yule Walker equation which allows minimizing equation (11) is as follows,

$$C_M a_M = c_M \quad (12)$$

where c_M denotes autocorrelation based covariance matrix and is expressed as follows,

$$c(k) = \frac{1}{N} \sum_{s=1}^{N-k} x(s)x(s+k)$$

$$C_M = \begin{bmatrix} c(0) & c(1) & \cdots & c(M-2) & c(M-1) \\ c(1) & c(0) & c(1) & \cdots & c(M-2) \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ c(M-2) & \cdots & c(1) & c(0) & c(1) \\ c(M-1) & c(M-2) & \cdots & c(1) & c(0) \end{bmatrix}$$

$$c_M = (c(1), c(2), \cdots, c(M))^T$$

Solution of Yule Walker equation (12) can be expressed as $\hat{a}_M$, then,

$$\hat{\sigma}_M^2 = c(0) - c_M^T \hat{a}_M$$

At this time, the maximum logarithmic function of likelihood is expressed as follows,

$$\log f(x|\hat{\sigma}_M^2, \hat{a}_M) = N \log(2\pi\hat{\sigma}_M^2)^{-\frac{1}{2}} - \frac{N}{2}$$

E. Time Series of Multi-Variables Prediction

The aforementioned prediction method can be used for prediction of time series of multi-variables,

$$\mathbf{X}_n = (x_n(1), \cdots, x_n(k))^T$$

$$E[\mathbf{U}(n)] = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$E\left[\mathbf{U}(n)\mathbf{U}(n)^T\right] = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1k} \\ \vdots & \ddots & \vdots \\ \sigma_{k1} & \cdots & \sigma_{kk} \end{bmatrix} = W$$

$$E\left[\mathbf{U}(n)\mathbf{U}(n)^T\right] = [0] \quad (n \neq m)$$
$$E\left[\mathbf{U}(n)\mathbf{X}(n-m)^T\right] = [0] \quad (m \geq 1)$$

Then the expectation is expressed as follows,

$$\begin{aligned} & E\left[\mathbf{U}(n)\mathbf{X}(n)^T\right] \\ = & E\left[\mathbf{U}(n)\{\sum_{m=1}^{M} \mathbf{X}(n-m)^T A(m)^T + \mathbf{U}(n)^T\}\right] \\ = & \sum_{m=1}^{M} E\left[\mathbf{U}(n)\mathbf{X}(n-m)^T\right] A(m)^T + E\left[\mathbf{U}(n)\mathbf{U}(n)^T\right] \\ = & E\left[\mathbf{U}(n)\mathbf{U}(n)^T\right] \\ = & W \end{aligned}$$

On the other hands, autocorrelation is expressed as follows,

$$R(m) = E\left[\mathbf{X}(n)\mathbf{X}(n-m)^T\right] = \begin{bmatrix} r_{11}(m) & \cdots & r_{1k}(m) \\ \vdots & \ddots & \vdots \\ r_{k1}(m) & \cdots & r_{kk} \end{bmatrix}$$

$$r_{ij}(m) = E[x_i(n)x_j(n-m)]$$

There is the following relation between the expectation and the autocorrelation,

$$\begin{aligned} R(-m) &= E\left[\mathbf{X}(n)\mathbf{X}(n+m)^T\right] \\ &= E\left[\mathbf{X}(n-m)\mathbf{X}(n)^T\right] \\ &= E\left[\mathbf{X}(n)\mathbf{X}(n-m)^T\right]^T \\ &= R(m)^T \end{aligned}$$

$$\begin{aligned} R(m) &= E\left[\left\{\sum_{j-1}^{M} A(j)\mathbf{X}(n-j) + \mathbf{U}(n)\right\} \mathbf{X}(n-m)^T\right] \\ &= \sum_{j=1}^{M} A(j) E\left[\mathbf{X}(n-j)\mathbf{X}(n-m)^T\right] + E\left[\mathbf{U}(n)\mathbf{X}(n-m)^T\right] \end{aligned}$$

Then

$$R(0) = \sum_{j=1}^{M} A(j)R(-j) + W$$
$$R(m) = \sum_{j=1}^{M} A(j)R(m-j) \quad (m = 1, 2, \cdots)$$

If the time series of multi variables are stationery, then

$$\begin{aligned} R(m) &= E\left[\mathbf{X}(n)\mathbf{X}(n-m)^T\right] \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbf{X}(n)\mathbf{X}(n-m)^T \end{aligned}$$

Then covariance matrix is expressed a s follows,

$$C(m) = \frac{1}{N} \sum_{n=m+1}^{N} \mathbf{X}(n)\mathbf{X}(n-m)^T$$
$$C_{ij}(m) = \frac{1}{N} \sum_{n=m+1}^{N} x_i(n)x_j(n-m)^T$$

[3] http://ja.wikipedia.org/wiki/%E8%87%AA%E5%B7%B1%E5%9B%9E%E5%B8%B0%E7%A7%BB%E5%8B%95%E5%B9%B3%E5%9D%87%E3%83%A2%E3%83%87%E3%83%AB

Thus autoregressive model of $\hat{A}(j)$ can be determined as follows,

$$C(m) = \sum_{j=1}^{M} A(j)C(m-j)$$

$$\hat{V}_m = C(0) - \sum_{j=1}^{M} \hat{A}(j)C(-j)$$

In order to determine the most appropriate order (or the number of possible predictions) of the autoregressive model, Akaike's Information Criteria: AIC is introduced,

AIC=-2L+2P

where L denotes the maximum logarithmic function of likelihood while P denotes the number of adjustable parameters. Therefore,

$$AIC(m) = Nk\log 2\pi + N\log|\hat{V}_m| + Nk + 2mk^2 + k(k+1)$$

where

$$L = -\frac{N}{2}(k\log 2\pi + \log|\hat{V}_m| + k)$$

$$P = mk^2 + \frac{k(k+1)}{2}$$

Remove the terms which are not related to the order of m,

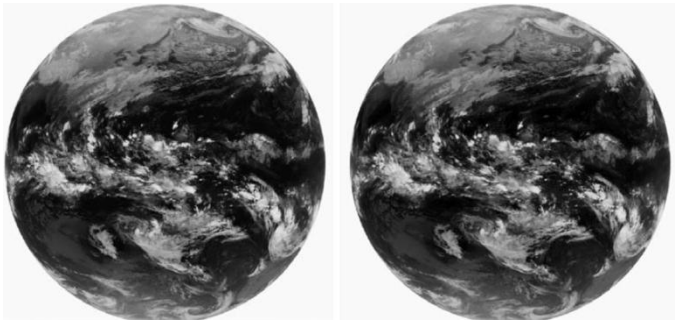
$$AIC^* = N\log|\hat{V}_m| + 2mk^2 \qquad (13)$$

is used to be used usually.

III. EXPERIEMNTS

A. Time Series of Imagery Data Used

Figure 1 shows experimental time series of imagery data of Visible/Infrared Spin Scan-Radiometer: VISSR onboard Geostationary Meteorological Satellite: GMS acquired time series of imagery data.



(a)Infrared 1 (b)Infrared 2

Figure 1 VISSR/GMS of time series of imagery data which was acquired at 16:25 on 12 December 2000.

VISSR/GMS has four channels, 0.55-0.9, 10.5-11.5, 11.5-12.5, 6.5-7.0 μ m, Infrared 1 denotes 10.5-11.5 while infrared 2 denotes 11.5-12.5, respectively.

VISSR/GMS imagery data can be acquired every one hour. From the VISSR/GMS images, 100 by 100 pixels of small portion of images are extracted. Portion of images of the same area of South East China sea area are extracted as shown in Figure 2.

Figure 2 Time series of VISSR/GMS images of infrared 1(Top) and 2(Bottom) channels of data used for experiments

B. AIC

41 of time series of VISSR/GMS infrared 1 and 2 channels of imagery data are used for prediction. AIC is estimated with the experimental data in the real world space and in the eigen space. AIC order in real space is 3, 5, 10 while that in eigen space is 2, 5, 10. Therefore, 10 of time series of imagery data of VISSR/GMS are tried to predict. Although autoregressive model parameters are estimated with 31 of imagery data, the imagery data which are acquired 10 hour later is available

AIC is calculated for each order using covariance matrix in accordance with equation (13). The results are shown in Table 1.

TABLE I. AIC CALCULATED WITH COVARIANCE METRIX

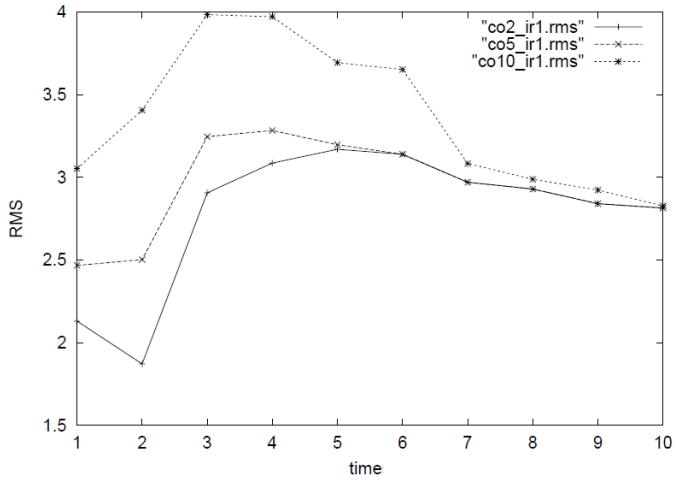
		Eigen space	Autoregressive
IR1	3	132.9	169.5
	5	139.3	174.3
	10	160.2	200.1
IR2	3	119.3	205.3
	5	127.8	210.8
	10	143.6	228.4
Two Variables	3	265.7	374.9
	5	252.1	361.4
	10	265.6	389.4

In the table, IR1 and IR2 denotes infrared 1 and 2 VISSR/GMS channels of imagery data while two variables means the case of prediction by using IR1 and IR2 together. AIC are different between eigen space and real space (the conventional autoregressive model). Therefore, both AICs are calculated. In general, smaller AIC implies better prediction accuracy.

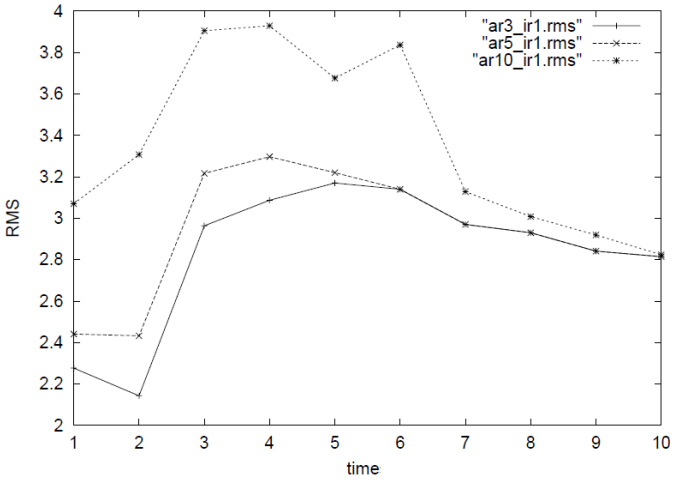
C. RMS Error

It is possible to evaluate Root Mean Square: RMS error between the ten of original images and the predicted images of the corresponding time. The estimated RMS errors are shown in Figure 3. Figure 3 shows RMS error for the single variable cases.

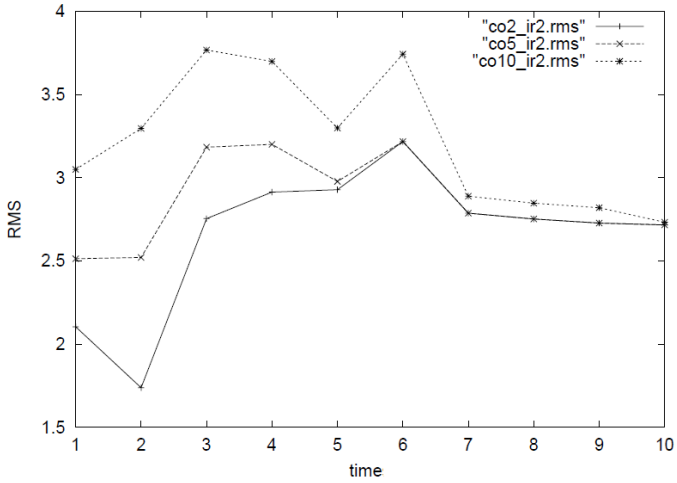
Figure 4 shows the experimental results of RMS error for the two variables cases.



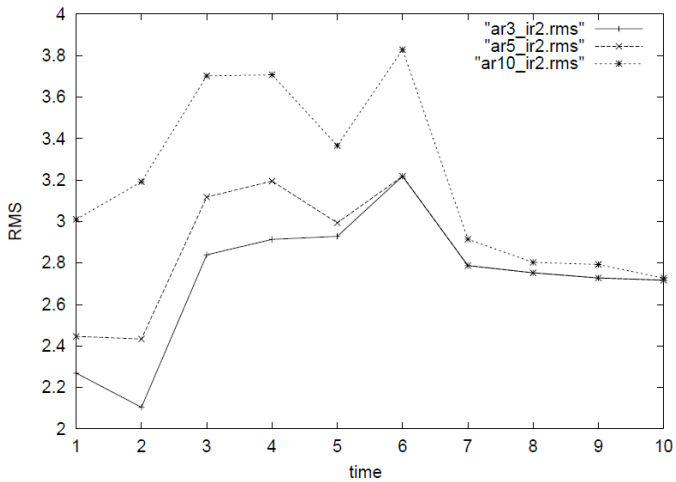
(a)Eigen space for IR1



(b)Real space for IR1

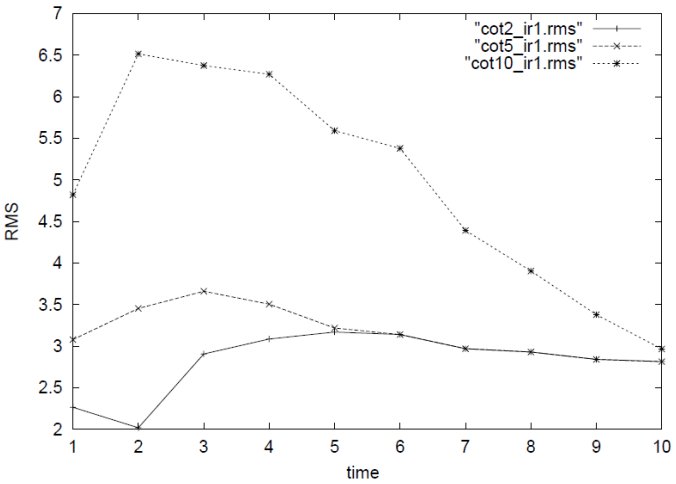


(c)Eigen space for IR2

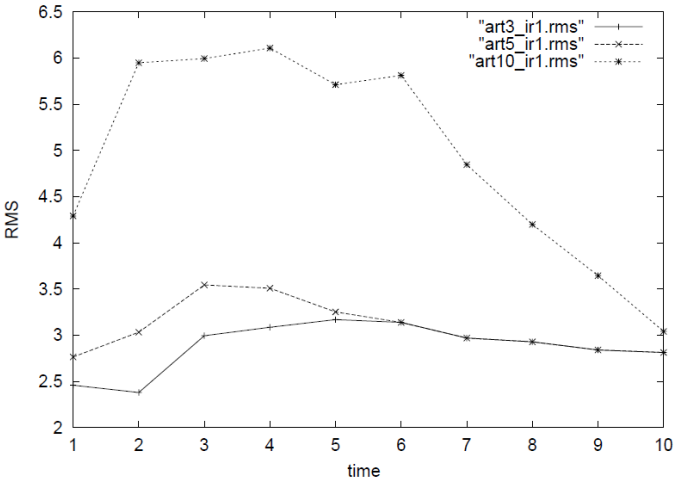


(d)Real space for IR2

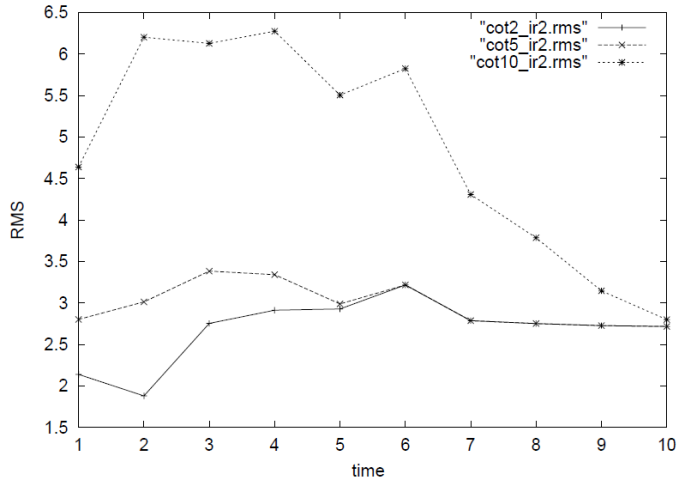
Figure 3 RMS error evaluated with VISSR/GMS time series of images for 10 unit time later data predictions based on the autoregressive models represented in real and eigen spaces for single variable cases.



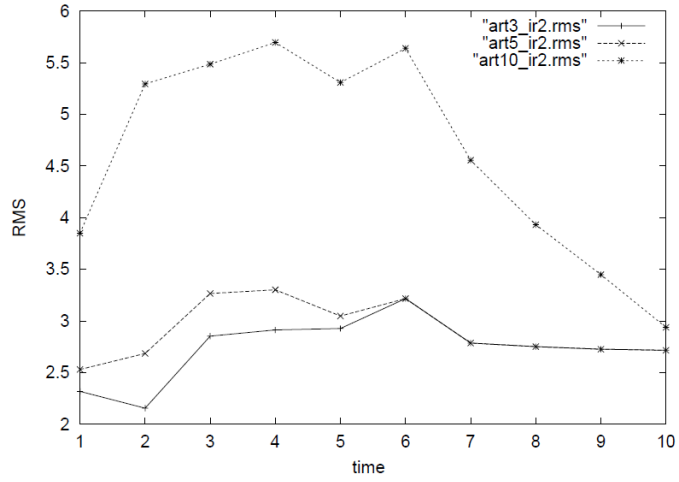
(a)Eigen space for IR1



(b)Real space for IR2



(c)Eigen space for IR1

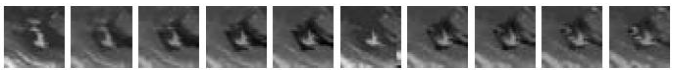


(d)Real space for IR2

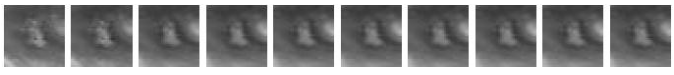
Figure 4 RMS error evaluated with VISSR/GMS time series of images for 10 unit time later data predictions based on the autoregressive models represented in real and eigen spaces for two variable cases.

D. Original and Predicted Images and RMS Error for IR1

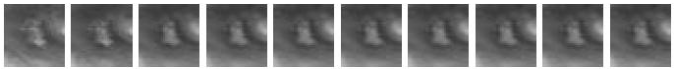
Figure 5 (b) shows the predicted images for the conventional real space based method while Figure 5 (c) shows those for the proposed eigen space based method.



(a)Original images



(b)Real space based



(c)Eigen space based

Figure 5 Original and predicted images based on the conventional real space and the proposed eigen space based methods

Figure 5 (a) shows the 10 hours of time series of the original portion of VISSR/GMS images while Figure 5 (b) shows the predicted images at the corresponding time for Figure 5 (a). Also Figure 5 (c) shows the predicted images at the corresponding time.

Prediction performances in terms of RMS error between the original and the predicted images for the conventional and the proposed methods are summarized in Table 2. Table 2 also shows improvement of the proposed method in comparison to the conventional real space based method. The RMS errors for the first to third images show some improvements of prediction accuracy in comparison between the conventional and the proposed method while the RMS errors for fourth or later images show no improvement. Then the proposed prediction method is effective for first to third time periods and is not effective for further time period.

As shown in Figure 6, RMS errors for real space (ar3_ir1.rms) and eigen space (co2_ir1.rms) are getting close each other

TABLE II. RMS ERROR BETWEEN THE ORIGINAL AND THE PREDICTED IMAGES BASED ON REAL SPACE AND EIGEN SPACE BASED METHODS FOR IR1

n-th image	Eigen space	Real space	Prediction improvement
1	2.131	2.276	6.4
2	1.871	2.143	12.7
3	2.907	2.963	1.9
4	3.086	3.086	0
5	3.17	3.17	0
6	3.139	3.139	0
7	2.97	2.97	0
8	2.929	2.929	0
9	2.841	2.841	0
10	2.814	2.814	0

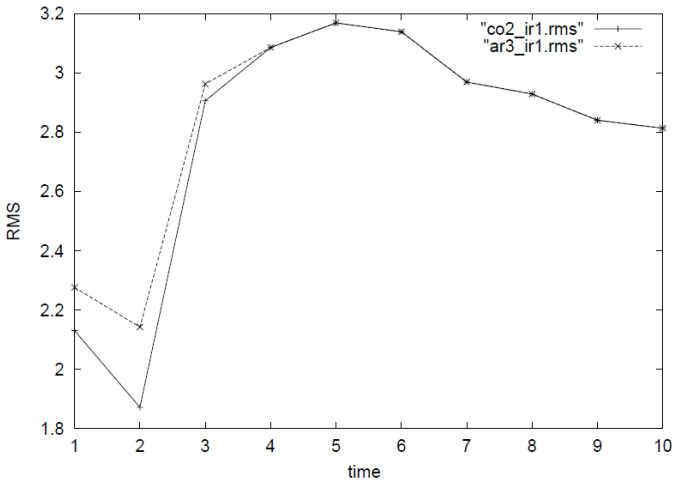
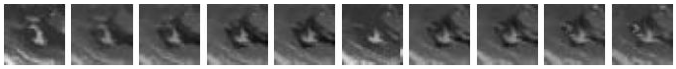


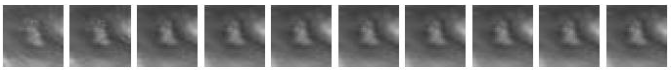
Figure 6 RMS errors for real space (ar3_ir1.rms) and eigen space (co2_ir1.rms)

E. Original and Predicted Images and RMS Error for IR2

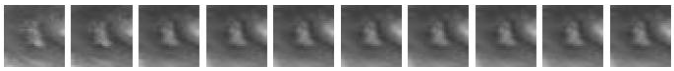
Figure 7 (b) shows the predicted images for the conventional real space based method while Figure 7 (c) shows those for the proposed eigen space based method.



(a)Original images



(b)Real space based



(c)Eigen space based

Figure 7 Original and predicted images based on the conventional real space and the proposed eigen space based methods

Figure 7 (a) shows the 10 hours of time series of the original portion of VISSR/GMS images while Figure 7 (b) shows the predicted images at the corresponding time for Figure 7 (a). Also Figure 7 (c) shows the predicted images at the corresponding time.

Prediction performances in terms of RMS error between the original and the predicted images for the conventional and the proposed methods are summarized in Table 3. Table 3 also shows improvement of the proposed method in comparison to the conventional real space based method. The RMS errors for the first to third images show some improvements of prediction accuracy in comparison between the conventional and the proposed method while the RMS errors for fourth or later images show no improvement. Then the proposed prediction method is effective for first to third time periods and is not effective for further time period.

As shown in Figure 8, RMS errors for real space (ar3_ir2.rms) and eigen space (co2_ir2.rms) are getting close each other

TABLE III. RMS ERROR BETWEEN THE ORIGINAL AND THE PREDICTED IMAGES BASED ON REAL SPACE AND EIGEN SPACE BASED METHODS FOR IR2

n-th image	Eigen space	Real space	Prediction improvement
1	2.104	2.269	7.3
2	1.738	2.104	17.4
3	2.754	2.839	3
4	2.913	2.913	0
5	2.928	2.928	0
6	3.217	3.217	0
7	2.787	2.787	0
8	2.752	2.752	0
9	2.727	2.727	0
10	2.717	2.717	0

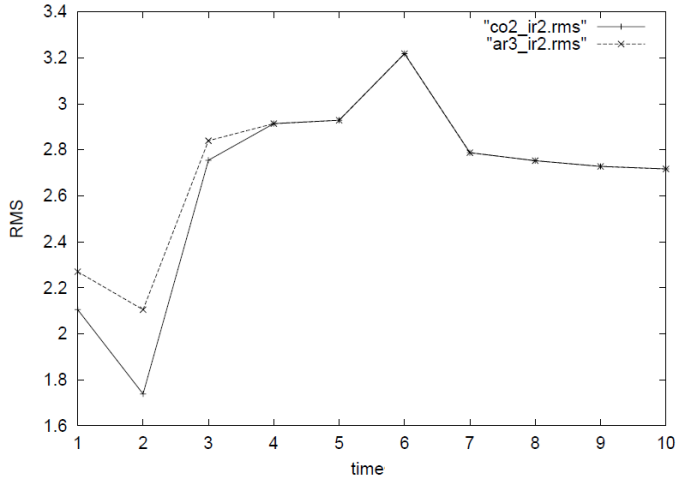
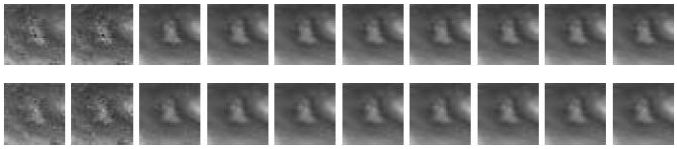


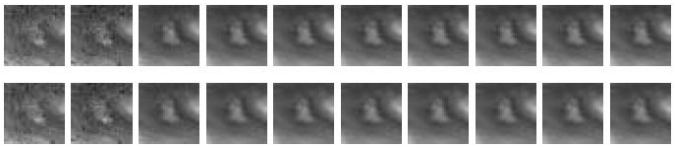
Figure 8 RMS errors for real space (ar3_ir2.rms) and eigen space (co2_ir12.rms)

F. Original and Predicted Images and RMS Error for Two Variables Cases

Figure 9 (a) shows the predicted images for the conventional real space based method while Figure 9 (b) shows those for the proposed eigen space based method. Figure 9 (a) shows the predicted images at the corresponding time for Figure 5 (a). Also Figure 9 (b) shows the predicted images at the corresponding time.



(a)Real space based



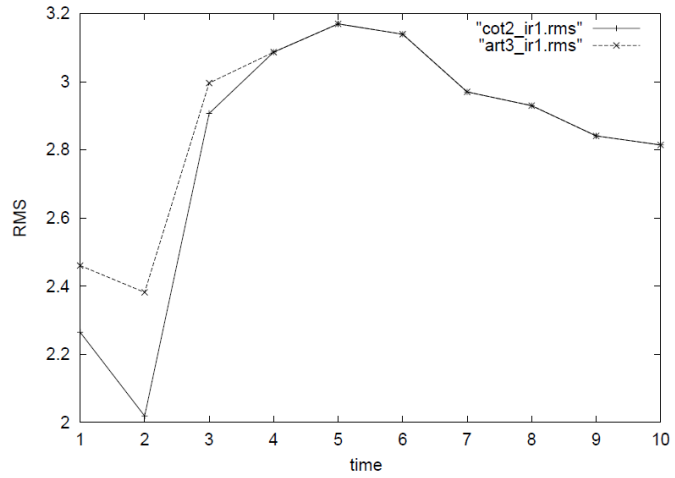
(b)Eigen space based

Figure 9 Original and predicted images based on the conventional real space and the proposed eigen space based methods

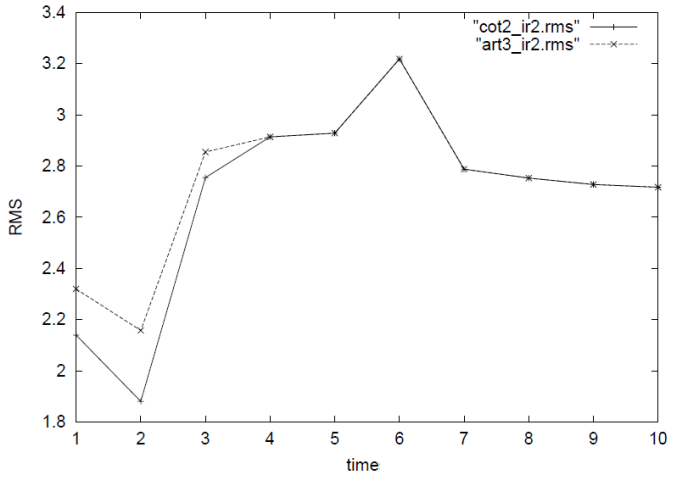
Prediction performances in terms of RMS error between the original and the predicted images for the conventional and the proposed methods are summarized in Table 4. Table 4 also shows improvement of the proposed method in comparison to the conventional real space based method. The RMS errors for the first to third images show some improvements of prediction accuracy in comparison between the conventional and the proposed method while the RMS errors for fourth or later images show no improvement. Then the proposed prediction method is effective for first to third time periods and is not effective for further time period. As shown in Figure 10, RMS errors for real space (ar3_ir2.rms) and eigen space (co2_ir2.rms) are getting close each other. This implies that the proposed time series analysis in eigen space is effective for the prediction within three recent unit times and is not effective for no further unit time.

TABLE IV. RMS ERROR BETWEEN THE ORIGINAL AND THE PREDICTED IMAGES BASED ON REAL SPACE AND EIGEN SPACE BASED METHODS FOR TWO VARIABLES CASES

	IR1			IR2		
n-th future	Eigen space	Real space	Prediction improvement	Eigen space	Real space	Prediction improvement
1	2.264	2.46	8	2.139	2.319	7.8
2	2.018	2.381	15.3	1.88	2.157	12.9
3	2.907	2.996	3	2.754	2.854	3.6
4	3.086	3.086	0	2.913	2.913	0
5	3.17	3.17	0	2.928	2.928	0
6	3.139	3.139	0	3.217	3.217	0
7	2.97	2.97	0	2.787	2.787	0
8	2.929	2.929	0	2.752	2.752	0
9	2.841	2.841	0	2.727	2.727	0
10	2.814	2.814	0	2.717	2.717	0



(a)IR1



(b)IR2

Figure 10 RMS errors for real space (ar3_ir2.rms) and eigen space (co2_ir12.rms)

IV. CONCLUSION

Prediction method for time series of imagery data on eigen space is proposed. Although the conventional prediction method is defined on the real world space and time domains, the proposed method is defined on eigen space. Prediction accuracy of the proposed method is supposed to be superior to the conventional methods. Through experiments with time series of satellite imagery data, validity of the proposed method is confirmed. The proposed time series analysis in eigen space is effective for the prediction within three recent unit times and is not effective for no further unit time

ACKNOWLEDGMENT

The author would like to tank Dr. Yasunori Terayama and Mr. Masaya Ishida for their effort to conduct experiments.

REFERENCES

[1] Bloomfield, P., Fourier analysis of time series: An introduction. New York: Wiley, 1976.

[2] Shumway, R. H., *Applied statistical time series analysis*. Englewood Cliffs, NJ: Prentice Hall, 1988.

[3] Gershenfeld, N., *The nature of mathematical modeling*. p.205-08, 1999.

[4] Boashash, B. (ed.), *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*, Elsevier Science, Oxford, 2003 ISBN ISBN 0-08-044335-4, 2003.

[5] Nikolić D, Muresan RC, Feng W, Singer W, Scaled correlation analysis: a better way to compute a cross-correlogram. *European Journal of Neuroscience*, pp. 1–21, doi:10.1111/j.1460-9568.2011.07987.x, 2012. http://www.danko-nikolic.com/wp-content/uploads/2012/03/Scaled-correlation-analysis.pdf

[6] K. Arai, Fundamental Theory and Algorithms on Image Processing, Gakujutsu-Tosho Publishing Co. Ltd., 1995.

AUTHORS PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science, and Technology of the University of Tokyo from 1974 to 1978 also was with National Space Development Agency of Japan (current JAXA) from 1979 to 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He was appointed professor at Department of Information Science, Saga University in 1990. He was appointed councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was also appointed councilor of Saga University from 2002 and 2003 followed by an executive councilor of the Remote Sensing Society of Japan for 2003 to 2005. He is an adjunct professor of University of Arizona, USA since 1998. He also was appointed vice chairman of the Commission "A" of ICSU/COSPAR in 2008. He wrote 30 books and published 332 journal papers

Image Prediction Method with Nonlinear Control Lines Derived from Kriging Method with Extracted Feature Points Based on Morphing

Kohei Arai
Graduate School of Science and Engineering
Saga University
Saga City, Japan

***Abstract*—Method for image prediction with nonlinear control lines which are derived from extracted feature points from the previously acquired imagery data based on Kriging method and morphing method is proposed. Through comparisons between the proposed method and the conventional linear interpolation and widely used Cubic Spline interpolation methods, it is found that the proposed method is superior to the conventional methods in terms of prediction accuracy.**

Keywords-Kriging; morphing; image prediction; interpolation; image feature extraction.

I. INTRODUCTION

Time series of images which are acquired with unequal interval can be converted to en equal interval of the time series of images. The proposed conversion method is based on morphing [1] method[1] utilizing feature points extracted from the time series of images. Control lines required for morphing method are defined by the extracted features. Then the time series of images at the designated time is interpolated with the adjacent images. In this process, nonlinear control lines are created in accordance with Kriging [2]-[15] method[2]. Through experiments with satellite imagery data, it is confirmed a validity of the proposed method in comparison to the conventional linear interpolations and widely used Cubic Spline[3] interpolation methods.

The following section describes the proposed method followed by some experiments. Then conclusion is described with some discussions.

II. PROPOSED METHOD

A. Process Flow of the Proposed Method

The proposed method is for conversion from unequally interval of time series images to equally interval of time series images. Therefore, some interpolations are necessary. There are many interpolation methods such as Cubic Spline interpolation (Appendix), Hermitian[4] interpolation, Bezier[5] interpolation, Lagrange[6] interpolation and so on. The proposed method is based on morphing method utilizing feature points extracted from the time series of images. Control lines required for morphing method are defined by the extracted features. Then the time series of images at the designated time is interpolated with the adjacent images. In this process, nonlinear control lines are created in accordance with Kriging method.

B. Morphing Method

Morphing method is "metamorphosing"[7] which allows interpolation from the objective shape to the other objective shapes smoothly. The basic idea of the morphing method is the following, shapes in an image represented on a rubber sheet is modified with pin movements elastically as shown in Figure 1. Figure 1 shows an example of the basic idea of morphing method.

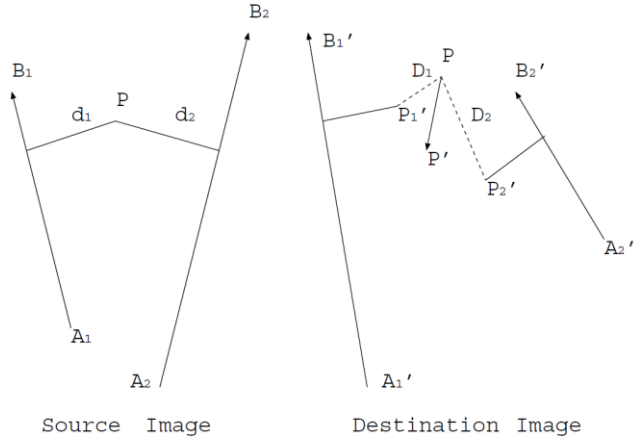


Figure 1 Basic idea of the morphing method

[1] http://www.enotes.com/topic/Godley_%26_Creme
[2] http://en.wikipedia.org/wiki/Kriging
[3] http://ja.wikipedia.org/wiki/%E3%82%B9%E3%83%97%E3%83%A9%E3%82%A4%E3%83%B3%E6%9B%B2%E7%B7%9A
[4] http://1st.geocities.jp/shift486909/program/hermite.html
[5] http://ja.wikipedia.org/wiki/%E3%83%99%E3%82%B8%E3%82%A8%E6%9B%B2%E7%B7%9A
[6] http://ja.wikipedia.org/wiki/%E3%83%A9%E3%82%B0%E3%83%A9%E3%83%B3%E3%82%B8%E3%83%A5%E8%A3%9C%E9%96%93
[7] http://www.thefreedictionary.com/metamorphosing

In the figure, the point "p" in the previous image moves to point "p"' with the weighted mean of the following two points, "p_1"' and "p_2"' where "p_1"' is orthogonally projected point of "p" against the line of A_1'~B_1' while "p_2"' is orthogonally projected point of "p" against the line of A_2'~B_2'.

C. Kriging Method

Kriging method allows estimation of future point locations with known point locations probabilistically. If only one known point location used, then future point location is exactly same location. If we know that a future point is situated at the center of two known points, the future points is situated at the point of median. Thus the future point, $\hat{y}(X_o)$ is estimated with equation (1) by using the known points, $y(X_j)$.

$$\hat{y}(X_o) = \sum_{j=1}^{n} \lambda_j y(X_j) \tag{1}$$

λ_j can be determined with the condition of equation (2) minimizing equation (3)

$$\sum_{j=1}^{n} \lambda_j = 1 \tag{2}$$

$$V_e = 2\sum_{j=1}^{n} \lambda_j \gamma(L_{jo}) - \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \gamma(L_{ij}) \tag{3}$$

where L_{jo} is defined as the distance between the future point,*0* and the known points, *j* while L_{ij} is defined as the distance in between the known points, *i* and *j*.

Meanwhile, estimation error covariance of semi-variogram, $\gamma(L)$ is defined in equation (4).

$$\gamma(L) = a(1 - exp(bL)) \tag{4}$$

where *a* and *b* are coefficients while *L* denotes distances. Therefore, if the coefficients are determined with the known points, then the distance is known. Thus λ_j can be determined results in estimation of future points.

III. Experiments

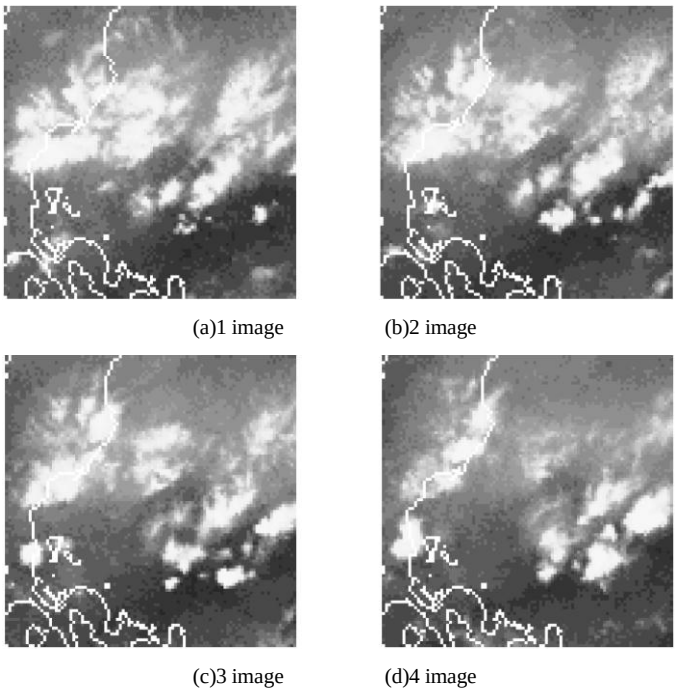
A. Imagery Data Used

Visible/Infrared Spin Scan-Radiometer: VISSR onboard Geostationary Meteorological Satellite: GMS acquired time series of imagery data are used for experiments. Figure 2 shows four time series of VISSR images of South East China Sea areas which were acquired in 1994. These are called image “A”, “B”, “C”, and “D”, respectively.

Although VISSR images are taken with one hour interval, interpolations are required for cloud tracking. For time being, clouds appear and disappear as well as move from one to the others as shown in Figure 2.

Also one hour interval would not be enough for cloud tracking, in particular, for the clouds which grow quite rapidly. Therefore, interpolations are necessary.

ppm file type of VISSR/GMS images are used for the experiments. Quantization bit of VISSR/GMS image is 8 bits (256 grey levels). 100 by 100 pixels are extracted from the original VISSR/GMS images.



(a)1 image (b)2 image

(c)3 image (d)4 image

Figure 2 Time series of VISSR/GMS images used

B. Experimental Procedure

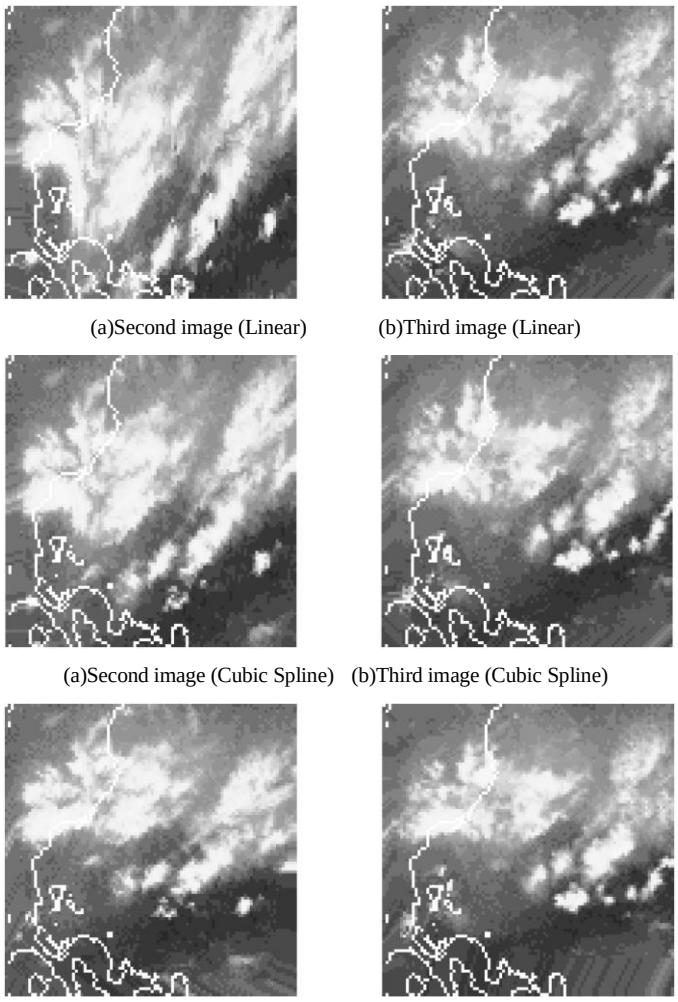
Assuming some of the time series of VISSR/GMS image is missing, interpolations are attempted for the supposed missing image by using the rest of given time series of images. Tie points are extracted from the two images which are acquired at the two closest times to the time at which the image is missing in accordance with the correlation between two images. By using the location of the extracted tie points, the location of the corresponding points on the missing image are estimated with the proposed and the other conventional interpolation methods, linear interpolation and Cubic Spline interpolation. Using the corresponding points, morphing method is applied to the two images which are acquired one previous and one after time to the missing image. Then the missing image is created as a result of the morphing.

Interpolation accuracy is evaluated with the difference between the interpolated and the original images which is supposed to be missing because the original image is known. In the evaluation process, two lines are estimated by using two corresponding points in between two previously acquired and just one after the image of which the image is missing. Evaluations of the difference are done for two cases of which the second and the third VISSR/GMS image are missing.

C. Experimental Results

Figure 3 (a) shows interpolated and predicted the second and the third images based on the conventional linear interpolation method while Figure 3 (b) shows those for the

Cubic Spline interpolation method, respectively. Meanwhile, Figure 3 (c) shows interpolated and predicted the second and .the third images based on the proposed method.



(a)Second image (Linear) (b)Third image (Linear)

(a)Second image (Cubic Spline) (b)Third image (Cubic Spline)

(a)Second image (Kriging) (b)Third image (Kriging)

Figure 3 Interpolated and predicted images

Root Mean Square Error: RMSE between the original and the interpolated and predicted images is evaluated. RMSE of the prediction of the second image "2" is 67.0, while that of the third image "3" is 53.5 for the linear interpolation method. Meanwhile, RMSE of the image "2" and "3" is 55.5, and 52.4, respectively for Cubic Spline interpolation method. On the other hands, RMSE of the image "2" and "3" is 54.1, and 51.9, respectively for the proposed Kriging utilized method. As the results, the proposed Kriging utilized method is superior to the conventional linear interpolation and Cubic Spline interpolation methods.

Actual locations of tie points are shown in Table 1. There are four tie points, "A", "B", "C", and "D" on each image. Using the tie points on images "1" and "3", tie points on image "2" is estimated while tie points on image "3" are estimated by using the tie points on images "2" and "4", respectively. The estimated tie point locations are shown in Table 2.

TABLE I. ACTUAL TIE POINT LOCATIONS: (X,Y) DENOTES X AND Y COORDINATES

	1	2	3	4
A	5,28	7,25	11,23	26,23
B	5,54	9,47	13,43	14,39
C	65,30	67,27	69,22	72,15
D	47,54	44,52	48,49	53,42

TABLE II. ESTIMATED TIE POINT LOCATIONS

A	2	3
Estimated location	A(7,25)	A(11,23)
Linear interpolation	A(7,26)	A(11,21)
Cubic Spline	A(7,26)	A(11,21)
Kriging	A(7,25)	A(11,25)
B	2	3
Estimated location	B(9,47)	B(13,43)
Linear interpolation	B(9,53)	B(13,41)
Cubic Spline	B(9,52)	B(13,42)
Kriging	B(9,45)	B(13,46)
C	2	3
Estimated location	C(67,27)	C(69,22)
Linear interpolation	C(67,26)	C(69,23)
Cubic Spline	C(67,26)	C(69,23)
Kriging	C(67,26)	C(69,24)
D	2	3
Estimated location	D(44,52)	D(48,49)
Linear interpolation	D(44,76)	D(48,53)
Cubic Spline	D(44,62)	D(48,52)
Kriging	D(44,48)	D(48,49)

Then RMS error between the actual image and the interpolated and predicted image by the proposed Kriging utilized method and the other interpolation method can be evaluated. Table 3 shows the results from the evaluation for the image size of 100 by 100 pixels, the original image size.

TABLE III. RMS ERROR OF THE ESTIMATED TIE POINT LOCATIONS

	RMS error(100x100)		
	Linear	Cubic Spline	Kriging
2	67	55.5	54.1
3	53.5	52.4	51.9

It shows that the proposed Kriging method utilized interpolation method is superior to the Linear and Cubic Spline interpolation methods. In more detailed RMS error is evaluated for 5 by 5 pixels of a small portion of images of which the tie points are centered of the portion of images for the cases to interpolate, predict images, "2" and "3".

Table 4 shows the result. Through comparisons between the proposed Kiriging utilized method is superior to the other conventional methods for all the cases. Meanwhile, RMS error between the original and interpolated images by using the actual tie point locations is evaluated for each 5 by 5 pixels portion images whose center is the tie point. Table 5 shows the results from the evaluation. Through the comparison between Table 4 and 5, it is found that the estimated tie points are very close to the actual points because RMS errors between Table 4 (Kriging) and 5 are almost same.

TABLE IV. RMS ERROR OF THE ESTIMATED TIE POINT LOCATIONS

	RMS error(5x5)		
2	Linear	Cubic Spline	Kriging
A	17.1	17.3	14
B	14.5	16.1	62.8
C	21.5	22.9	24
D	90.3	94.9	34.4
3	Linear	Cubic Spline	Kriging
A	23.1	23.1	16.5
B	17.5	16.8	21.2
C	40.2	40.2	32.1
D	57.6	44.8	19.8

TABLE V. RMS ERROR OF THE ESTIMATED TIE POINT LOCATIONS

5x5	2	3
A	13.4	15.7
B	37.4	15.6
C	23.4	44.3
D	25.2	20.1

IV. CONCLUSION

Method for image prediction with nonlinear control lines which are derived from extracted feature points from the previously acquired imagery data based on Kriging method and morphing method is proposed. Through comparisons between the proposed method and the conventional linear interpolation and widely used Cubic Spline interpolation methods, it is found that the proposed method is superior to the conventional methods in terms of prediction accuracy.

APPEXDIN: CUBIC SPLINE INTERPOLATION

n-th Spline function is defined as a n-th order polynomial function for a finite area and is a continuous function for all the defined areas which ensure continuously differentiable by n times. The third order spline function is defined as follows,

Assuming the n points,

$$(x_1, y_1), (x_2, y_2), ...(x_n, y_n)(a \le x_1 \le x_2 \le ..... \le x_n \le b)$$

the third order of spline function, $S_i(x)$ for the region, $[x_i, x_{i+1}]$ can be expressed as follows,

$$S_i(x) = a_i + b_i(x - x_i) + c_i(x - x_i)^2 + d_i(x - x_i)^3 (i = 1, 2, ..., n - 1)$$

From the condition of interpolation, the following equations have to be existed,

$$S_i(x_i) = y_i \qquad (i = 1, 2, ...n - 1)$$
$$S_i(x_{i+1}) = y_{i+1} \qquad (i = 1, 2, ...n - 1)$$

The first and the second derivatives have to be continuous. Therefore,

$$S_i^I(x_i) = S_{i-1}^I(x_i) \qquad (i = 2, 3, ...n - 1)$$
$$S_i^{II}(x_i) = S_{i-1}^{II}(x_i) \qquad (i = 2, 3, ...n - 1)$$

These are rewrite as follows,

$$S_i^I(x) = b_i + 2c_i(x - x_i) + 3d_i(x - x_i)^2 \qquad (i = 1, 2, ..., n - 1)$$
$$S_i^{II}(x) = 2c_i + 6d_i(x - x_i) \qquad (i = 1, 2, ..., n - 1)$$

If the following equation is assumed,

$$h_i = x_{i+1} - x_i \qquad (i = 1, 2, ..., n - 1)$$

then

$$a_i = y_i \qquad (i = 1, 2, ..., n - 1)$$
$$y_i + b_i h_i + c_i {h_i}^2 + d_i {h_i}^3 = y_{i+1} \qquad (i = 1, 2, ..., n - 1)$$
$$c_i + 3d_i h_i = c_{i+1} \qquad (i = 1, 2, ..., n - 2)$$

Thus

$$b_i = \frac{y_{i+1} - y_i}{h_i} - \frac{h_i}{3}(2c_{i+1} + c_i) \qquad (i = 1, 2, ..., n - 2)$$
$$d_i = \frac{c_{i+1} - c_i}{3h_i} \qquad (i = 1, 2, , ..., n - 1)$$
$$b_i = b_{i-1} + 2c_{i-1}h_{i-1} + 3d_{i-1}{h_{i-1}}^2 \qquad (i = 2, 3, ..., n - 1)$$

So that,

$$h_{i-1}c_{i-1} + 2(h_{i-1} + h_i)c_i + h_i c_{i+1}$$
$$= 3\left(\frac{y_{i+1} - y_i}{h_i} - \frac{y_i - y_{i-1}}{h_{i-1}}\right) \qquad (i = 2, 3, ..., n - 2)$$

In this equation, there aren-1ofunknownparameters,

$$C_i (i = 1, 2, ..., n - 1)$$

Also the number of the given equations is n-3. Therefore, other two conditional equations are required. $S(x)$ is essentially natural spline function. Therefore,

$$S_1''(x_1) = 0 \qquad S_{n-1}''(x_n) = 0$$
$$S_1'(x_1) = S_1' \qquad S_1'(x_n) = S_n'$$

Thus,

$$c_1 = 0 \qquad c_n = 0$$

And

$$2h_1c_1 + h_1c_2 = 3(\frac{y_2 - y_1}{h_1} - s_1')$$
$$h_{n-1}c_{n-1} + 2h_{n-1}c_n = 3(s_n' - \frac{y_n - y_{n-1}}{h_{n-1}})$$

Thus all the required coefficients for spline function are obtained.

ACKNOWLEDGMENT

The author would like to thank Mr. Shunji Tanaka for his experimental contributions through this study.

REFERENCES

[1] Scot Anderson, (2006) Triage, Fest, Alexander Verlag 978-3828600348 , p.286,

[2] Hanefi Bayraktar and F. Sezer. Turalioglu (2005) "A Kriging-based approach for locating a sampling site—in the assessment of air quality, SERRA, 19 (4), 301-305 doi:10.1007/s00477-005-0234-8

[3] Chiles, J.-P. and P. Delfiner (1999) Geostatistics, Modeling Spatial Uncertainty, Wiley Series in Probability and statistics.

[4] Zimmerman, D.A. et al. (1998) "A comparison of seven geostatistically based inverse approaches to estimate transmissivies for modelling advective transport by groundwater flow", Water Resources Research, 34 (6), 1273-1413

[5] Tonkin M.J. Larson (2002) "Kriging Water Levels with a Regional-Linear and Point Logarithmic Drift", Ground Water, 33 (1), 338-353,

[6] Journel, A.G. and C.J. Huijbregts (1978) Mining Geostatistics, Academic Press London

[7] Andrew Richmond (2003) "Financially Efficient Ore Selection Incorporating Grade Uncertainty", Mathematical Geology, 35 (2), 195-215

[8] Goovaerts (1997) Geostatistics for natural resource evaluation, OUP. ISBN 0-19-511538-4

[9] X. Emery (2005) "Simple and Ordinary Kriging Multigaussian Kriging for Estimating recoverarble Reserves", Mathematical Geology, 37 (3), 295-31)

[10] A. Stein, F. van der Meer, B. Gorte (Eds.) (2002) Spatial Statistics for remote sensing. Springer. ISBN 0-7923-5978-X

[11] Barris, J. (2008) An expert system for appraisal by the method of comparison';'. PhD Thesis, UPC, Barcelona

[12] Sacks, J. and Welch, W.J. and Mitchell, T.J. and Wynn, H.P. (1989). Design and Analysis of Computer Experiments. **4**. Statistical Science. pp. 409–435.

[13] Strano, M. (2008). "A technique for FEM optimization under reliability constraint of process variables in sheet metal forming". International Journal of Material Forming **1**: 13–20. doi:10.1007/s12289-008-0001-8. edit

[14] Grace Wahba (1990). Spline Models for Observational Data. **59**. SIAM. p. 162.

[15] Williams, Christopher K.I. (1998). "Prediction with Gaussian processes: From linear regression to linear prediction and beyond". In M. I. Jordan. Learning in graphical models. MIT Press. pp. 599–612.

AUTHORS PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science, and Technology of the University of Tokyo from 1974 to 1978 also was with National Space Development Agency of Japan (current JAXA) from 1979 to 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He was appointed professor at Department of Information Science, Saga University in 1990. He was appointed councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was also appointed councilor of Saga University from 2002 and 2003 followed by an executive councilor of the Remote Sensing Society of Japan for 2003 to 2005. He is an adjunct professor of University of Arizona, USA since 1998. He also was appointed vice chairman of the Commission “A” of ICSU/COSPAR in 2008. He wrote 30 books and published 332 journal papers

Method for object motion characteristic estimation based on wavelet Multi-Resolution Analysis: MRA

Kohei Arai
Graduate School of Science and Engineering
Saga University
Saga City, Japan

***Abstract*—Method for object motion characteristic estimation based on wavelet Multi-Resolution Analysis: MRA is proposed. With moving pictures, the motion characteristics, direction of translation, roll/pitch/yaw rotations can be estimated by MRA with an appropriate support length of the base function of wavelet. Through simulation study, method for determination of the appropriate support length of Daubechies base function is clarified. Also it is found that the proposed method for object motion characteristics estimation is validated.**

Keywords-object motion characteristic; MRA; wavelet.

I. INTRODUCTION

There are some conventional methods for object detections and object motion characteristics estimations. Optical flow, template matching is well known as the method for object motion characteristic estimations. On the other hands, the method for motion characteristic estimation based on wavelet MRA is also widely used [1]-[9]. Wavelet base function is defined with support length. The most appropriate support length depends on the motion characteristic [10],[11]. Therefore, it is necessary to determine an appropriate support length for estimation of motion characteristics. Meanwhile, it is required to estimate motion characteristics such as translation vector, speed and direction, rotation, direction and speed (rotation angle). 3D object motion characteristics is estimated based on wavelet analysis as well [12].

The method proposed here is based on wavelet MRA with an appropriate support length which is determined by the proposed method and allows estimation of motion characteristics with moving picture. Through simulation study, it is confirmed that the method for appropriate support length is valid together with the method for motion characteristic estimations.

The following section describes the proposed method followed by the simulation study.

II. PROPOSED METHOD

A. Process Flow of the Proposed Method

Figure 1 shows Process flow of the proposed motion characteristic estimations based on wavelet MRA with appropriate support length of base function which is determined by the proposed method. Acquired moving picture is essentially represented in three dimensional space. From the moving picture, the most appropriate support length is determined based on the proposed method which is described in the next sub-section. Then 3D wavelet MRA is applied to the acquired moving picture. As the results from the MRA, motion characteristics are estimated.

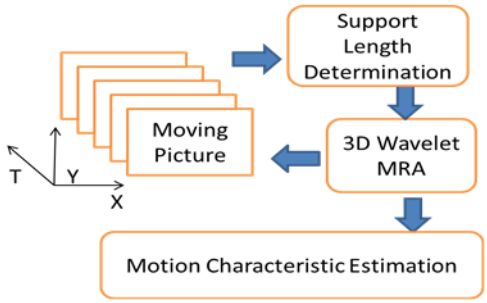


Figure 1 Process flow of the proposed motion characteristic estimations based on wavelet MRA with appropriate support length of base function which is determined by the proposed method

B. Wavelet Multi-Resolution Analysis: MRA

Two dimensional wavelet transformations is defined by equation (1).

$$F = [C_n[C_m f_{xy}]^t] \quad (1)$$

where C_m denotes another wavelet transformation matrix while f_{xy} denotes two dimensional data of image. As the result, $F=(LL_1, LH_1, HL_1, HH_1)$ with four frequency components in x and y directions, LL_1, LH_1, HL_1, HH_1. C_n can be determined with the method which is described in the second paragraph of this sub-section. Therefore, $C_n\ C_n^t=I$. Then f is converted to $F_1=(L_1,\ H_1)$, $F_2=C_nL_1=(L_2,\ H_2)$, $F_3=C_nL_2=(L_3,\ H_3)$, and $F_m=C_nL_{m-1}=(L_m, H_m)$.

This process is referred to "Decomposition". Also f is reconstructed as $C_n^{-1}F_m=C_n^{-1}(L_m, H_m)=L_{m-1},.....,C_n^{-1}F_2=L_1$, $C_n^{-1}F_1=f$. This process is referred to "Reconstruction". There are some based functions such as Haar, Daubechies, etc. Through the preliminary simulation study with radar echo data, Daubechies base function is selected. Daubechies base function is defined as $\{\alpha_k\}$ satisfying the following three conditions,

$$\phi(x) = \sum_k \alpha_k \sqrt{2}\phi(2x-k) \quad (2)$$

$$\beta_k = (-1)^k \alpha_{1-k} \tag{3}$$

$$\varphi(x) = \sum_k \beta_k \sqrt{2}\phi(2x-k) \tag{4}$$

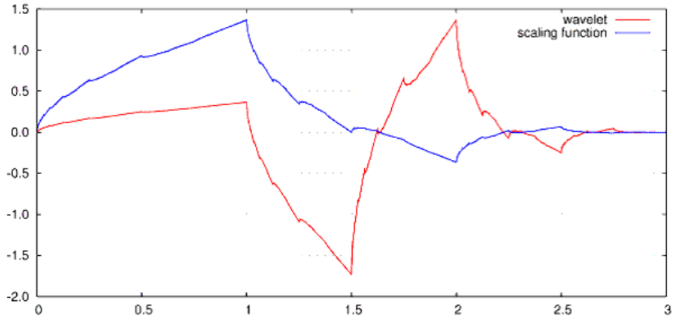
where equation (2) is referred to scaling function while equation (4) is referred to wavelet function, respectively. There is two-scale relation between the scaling function and the wavelet function. Also, *k* denotes support length.

Example of the scaling function and wavelet function with support length of four is as follows,

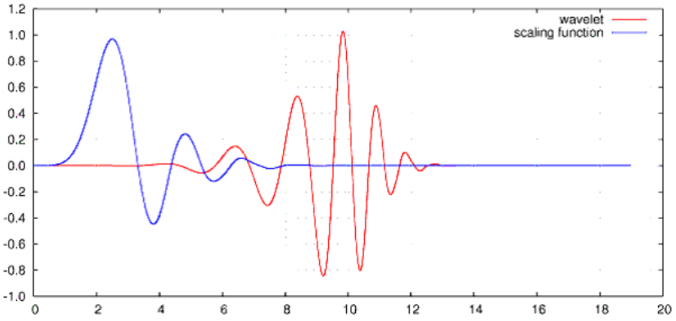
$$\phi(x) = \frac{1+\sqrt{3}}{4\sqrt{2}}\phi(2x-0) + \frac{3+\sqrt{3}}{4\sqrt{2}}\phi(2x-1) + \frac{3-\sqrt{3}}{4\sqrt{2}}\phi(2x-2) + \frac{1-\sqrt{3}}{4\sqrt{2}}\phi(2x-3)$$

$$\psi(x) = \frac{1-\sqrt{3}}{4\sqrt{2}}\psi(2x-0) - \frac{3-\sqrt{3}}{4\sqrt{2}}\psi(2x-1) + \frac{3+\sqrt{3}}{4\sqrt{2}}\psi(2x-2) - \frac{1+\sqrt{3}}{4\sqrt{2}}\psi(2x-3)$$

Both functions are shown in Figure 2.



(a)Support length=4



(b)Support length=20

Figure 2 Scaling and wavelet functions with the different support length.

Method for determination of Daubechies wavelet coefficients is as follows,

For instance, the 8th order[1] of Daubechies base function based C_n with the support length of two can be expressed with the equation (5).

$$C_8^{[2]}\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\\ \eta_5\\ \eta_6\\ \eta_7\\ \eta_8\end{bmatrix} = \begin{bmatrix} p_0 & p_1 & & & & & & \\ q_0 & q_1 & & & & & & \\ & & p_0 & p_1 & & & & \\ & & q_0 & q_1 & & & & \\ & & & & p_0 & p_1 & & \\ & & & & q_0 & q_1 & & \\ & & & & & & p_0 & p_1 \\ & & & & & & q_0 & q_1 \end{bmatrix}\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\\ \eta_5\\ \eta_6\\ \eta_7\\ \eta_8\end{bmatrix}$$

$$= \begin{bmatrix} p_0x_1 + p_1x_2 \\ q_0\eta_1 + q_1\eta_2 \\ p_0\eta_3 + p_1\eta_4 \\ q_0\eta_3 + q_1\eta_4 \\ p_0\eta_5 + p_1\eta_6 \\ q_0\eta_5 + q_1\eta_6 \\ p_0\eta_7 + p_1\eta_8 \\ q_0\eta_7 + q_1\eta_8 \end{bmatrix} \tag{5}$$

where [n] denotes support length. The 8th order of C_n with the support length of four is also expressed with the equation (6).

$$C_8^{[4]}\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\\ \eta_5\\ \eta_6\\ \eta_7\\ \eta_8\end{bmatrix} = \begin{bmatrix} p_0 & p_1 & p_2 & p_3 & & & & \\ q_0 & q_1 & q_2 & q_3 & & & & \\ & & p_0 & p_1 & p_2 & p_3 & & \\ & & q_0 & q_1 & q_2 & q_3 & & \\ & & & & p_0 & p_1 & p_2 & p_3 \\ & & & & q_0 & q_1 & q_2 & q_3 \\ p_2 & p_3 & & & & & p_0 & p_1 \\ q_2 & q_3 & & & & & q_0 & q_1 \end{bmatrix}\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\\ \eta_5\\ \eta_6\\ \eta_7\\ \eta_8\end{bmatrix}$$

$$= \begin{bmatrix} p_0\eta_1 + p_1\eta_2 + p_2\eta_3 + p_3\eta_4 \\ q_0\eta_1 + q_1\eta_2 + q_2\eta_3 + q_3\eta_4 \\ p_0\eta_3 + p_1\eta_4 + p_2\eta_5 + p_3\eta_6 \\ q_0\eta_3 + q_1\eta_4 + q_2\eta_5 + q_3\eta_6 \\ p_0\eta_5 + p_1\eta_6 + p_2\eta_7 + p_3\eta_8 \\ q_0\eta_5 + q_1\eta_6 + q_2\eta_7 + q_3\eta_8 \\ p_0\eta_7 + p_1\eta_8 + p_2\eta_1 + p_3\eta_2 \\ q_0\eta_7 + q_1\eta_8 + q_2\eta_1 + q_3\eta_2 \end{bmatrix} \tag{6}$$

p_i and q_i in the equations (5) and (6) is also expressed with equation (7) and (8), respectively.

$$\begin{aligned} \left(C_n^{[2]}\right)^T C_n^{[2]} &= I_n \\ p_0 + p_1 &= \sqrt{2} \\ q_0 &= p_1 \\ q_1 &= -p_0 \\ 0^0 q_0 + 1^0 q_1 &= 0 \end{aligned} \tag{7}$$

[1] Scalar data consist of eight numerical data, η_1 to η_8 is assumed.

$$\begin{aligned} \left(C_n^{[4]}\right)^T C_n^{[4]} &= I_n \\ p_0 + p_1 + p_2 + p_3 &= \sqrt{2} \\ q_0 &= p_3 \\ q_1 &= -p_2 \\ q_2 &= p_1 \\ q_3 &= -p_0 \\ 0^0 q_0 + 1^0 q_1 + 2^0 q_2 + 3^0 q_3 &= 0 \\ 0^1 q_0 + 1^1 q_1 + 2^1 q_2 + 3^1 q_3 &= 0 \end{aligned} \tag{8}$$

These equations can be expanded to the general support length of C_n as shown in equation (9).

$$\begin{aligned} \left(C_n^{[\sup]}\right)^T C_n^{[\sup]} &= I_n \\ \sum_{j=0}^{\sup-1} p_j &= \sqrt{2} \\ q_j &= (-1)^j p_{((\sup-1)-j)} \quad (j = 0,1,2,\ldots,(\sup-1)) \\ \sum_{j=0}^{\sup-1} j^r q_j &= 0 \quad \left(r = 0,1,2,\ldots,(\frac{\sup}{2}-1)\right) \end{aligned} \tag{9}$$

Therefore, the coefficients of the Daubechies base function can be determined from the solution of the equation (9).

C. Method for Most Appropriate Support Length Determinations

Figure 3 shows process flow of the proposed most appropriate support length determination method.

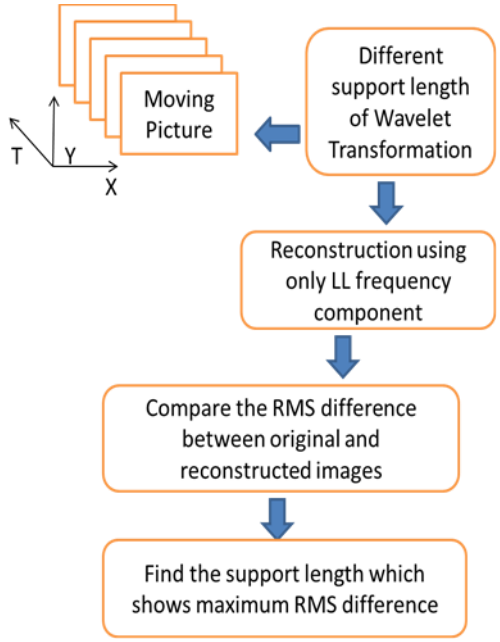


Figure 3 Method for the most appropriate support length determinations

Wavelet transformation with the different support length is applied to the original image. Then reconstruct the image with the decomposed images of LL component only. If the original image does not have any high frequency components, then LL component is totally equal to the original image. It is not always. Usually, most images have some extent of high frequency components results in the reconstructed image with LL component only is different from the original image obviously.

High frequency components are depends on the spatially and temporally changes in the original image. Therefore, if the Root Mean Square: RMS difference between the reconstructed and the original image is zero, then it may say that there is no high frequency component in the original image. Nevertheless the original image contains relatively large high frequency components, RMS difference is small then it may also say that the spatial and temporal changes cannot be extracted effectively. Consequently, it may say that the support length of which wavelet transformation with the different support length of base function is applied to the original image is the most appropriate if the RMS difference is maximum.

III. Simulation Study

A. Simulation Data Used

32 frames of time series of shuttlecock like simulation images are crated with PovRay of Computer Graphics software. Translations in vertical, horizontal, and slant directions, and the rotations in pitch, roll, and yaw directions are taken into account as motion characteristics. One of the examples of the moving picture of shuttlecock for the rotation in pitch direction is shown in Figure 4.

Figure 4 Example of the moving picture of shuttlecock for the rotation in pitch direction

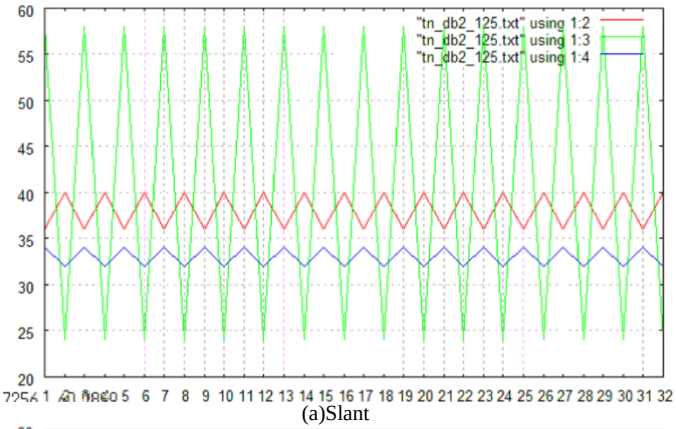
B. Simulation Procedure

2D wavelet transformation is applied to the simulation data of moving pictures. Then the reconstructed images are created with HL, LH, and HH components only. Thus the changed pixels or edge pixels are extracted. The changed pixels are counted after the thresholding (binarizing) with an appropriate threshold.

C. Simulation Results

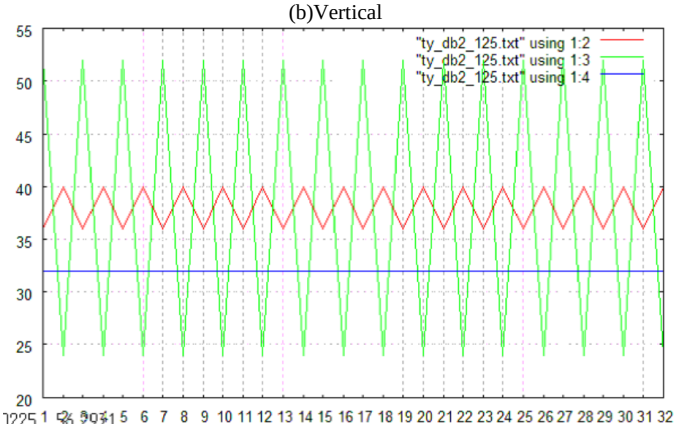
Figure 5 shows the number of changed pixels in the directions of vertical (Red), horizontal (Blue), and slant (Green), respectively when the shuttlecock moves in the directions of vertical, horizontal, and slant (45 degree). The resultant graphs are quite reasonable. In this case, the changed pixels are extracted based on the wavelet transformation with the Daubechies base function of support length of two. Therefore, it is possible to find the moving directions easily. On the other hands, Figure 6 shows the number of changed pixels when the shuttlecock moves in horizontal direction. In this case, the changed pixels are extracted based on the wavelet transformation with the Daubechies base function of the different support length, 2, 4, and 8.

Figure 6 (a), (b), and (c) shows the number of changed pixels extracted from the wavelet transformation with the Daubechies base function of the support length of 2, 4, and 8, respectively.



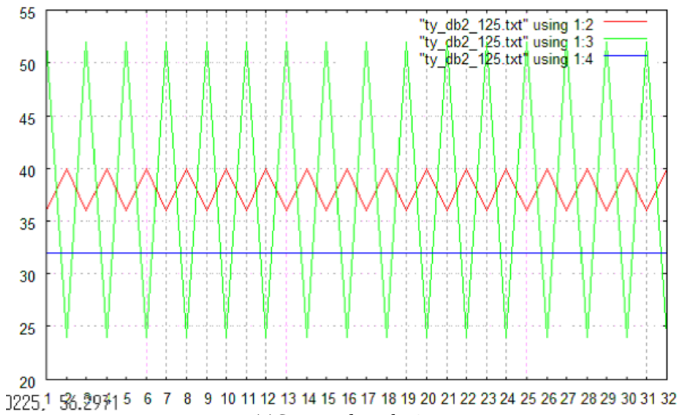
(a)Slant

(b)Vertical

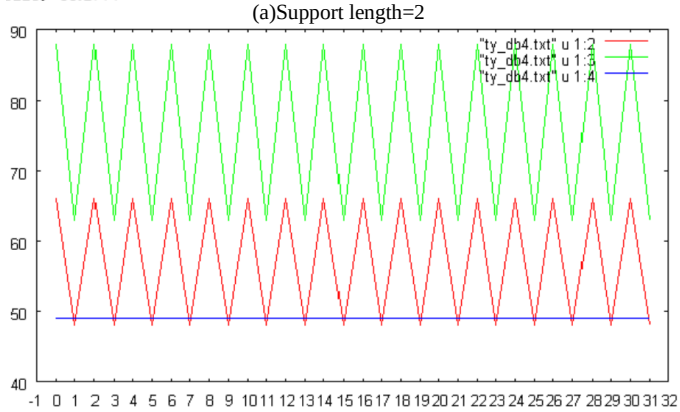


(c)Horizontal

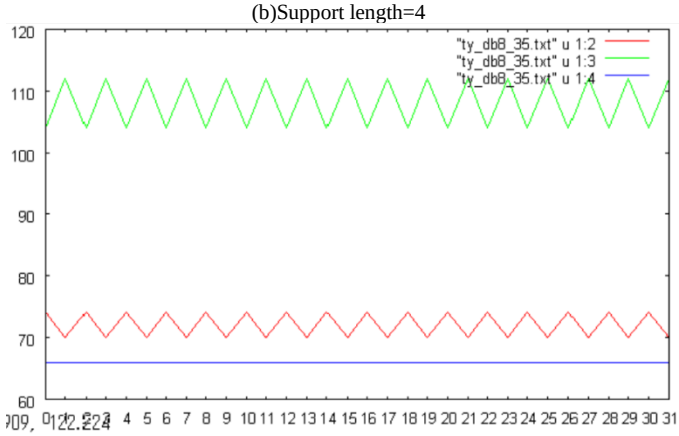
Figure 5 The number of changed pixels in the directions of vertical (Red), horizontal (Blue), and slant (Green), respectively when the shuttlecock moves in the directions of vertical, horizontal, and slant (45 degree). Changed pixels are extracted based on the wavelet transformation with support length of two.



(a)Support length=2



(b)Support length=4



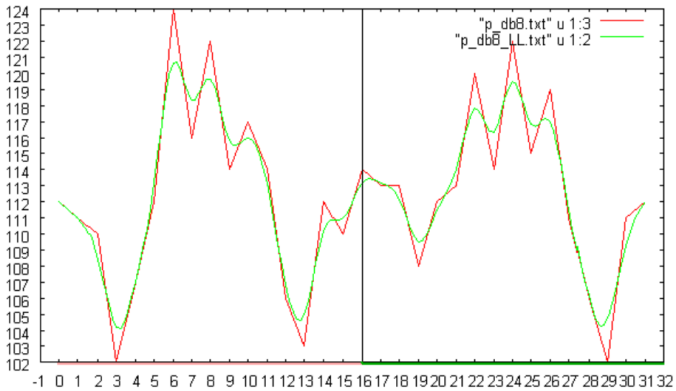
(c)Support length=8

Figure 6 The number of changed pixels when the shuttlecock moves in horizontal direction. In this case, the changed pixels are extracted based on the wavelet transformation with the Daubechies base function of the different support length, 2, 4, and 8.

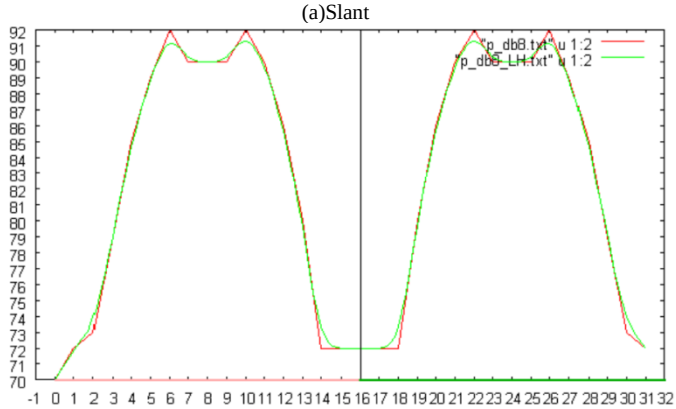
It is quite obvious that the extracted changed pixels depend on the support length as shown in Figure 6. Also, it is found that support length of 4 is the most appropriate.

Figure 7 shows the number of changed pixels in the directions of slant, vertical, and horizontal. In this case, shuttlecock rotates from 0 to 360 degrees in pitch direction.

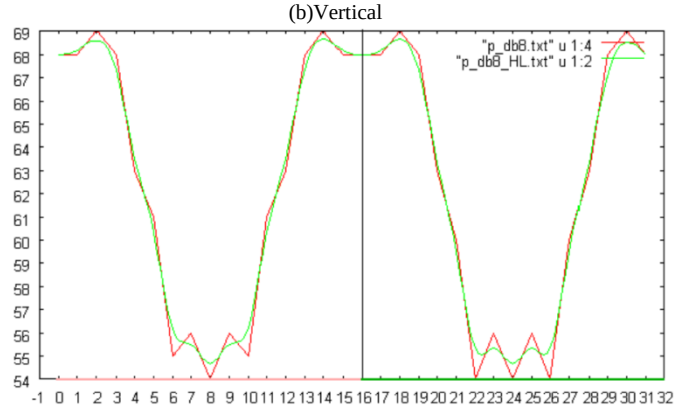
The changed pixels are extracted based on the wavelet transformation with the Daubechies base function of support length of 8. Meanwhile, Figure 8 (a), (b), and (c) shows the number of changed pixels in vertical direction, for the support length of 2, 4, and 8, respectively. Figure 9 (a), (b), and (c) shows the number of changed pixels when the shuttlecock is rotated by 180, 360, and 720 degrees, respectively.



(a)Slant

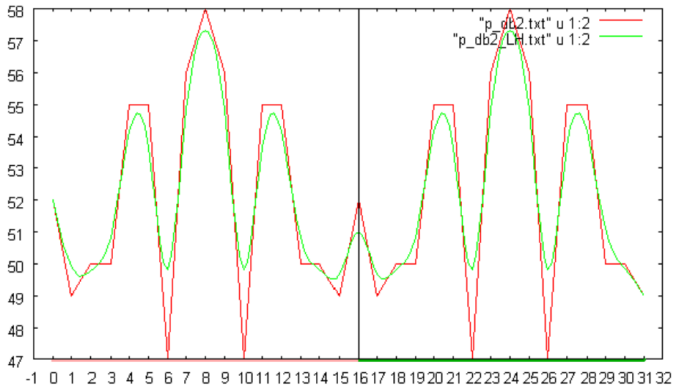


(b)Vertical

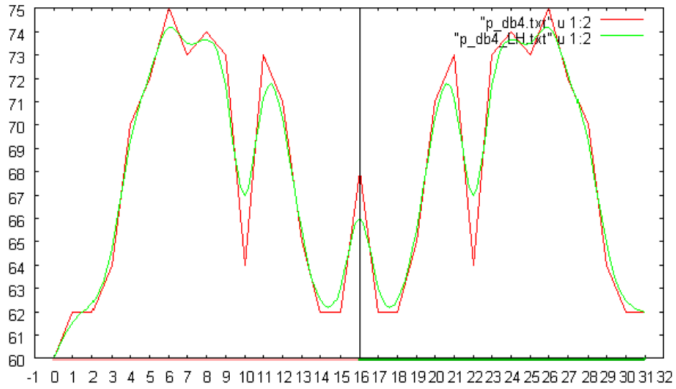


(c)Horizontal

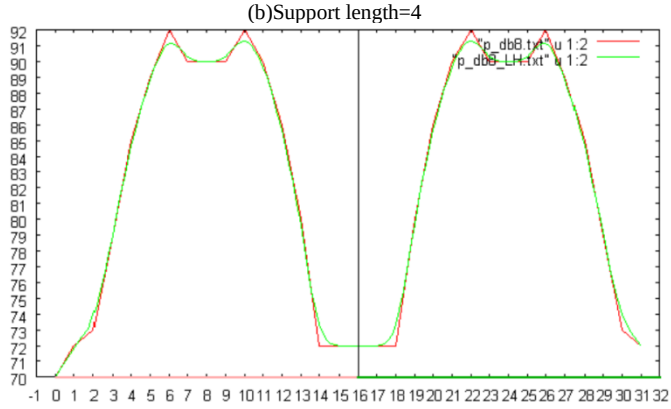
Figure 7 the number of changed pixels in the directions of slant, vertical, and horizontal. In this case, shuttlecock rotates from 0 to 360 degrees in pitch direction. The changed pixels are extracted based on the wavelet transformation with the Daubechies base function of support length of 8.



(a)Support length=2

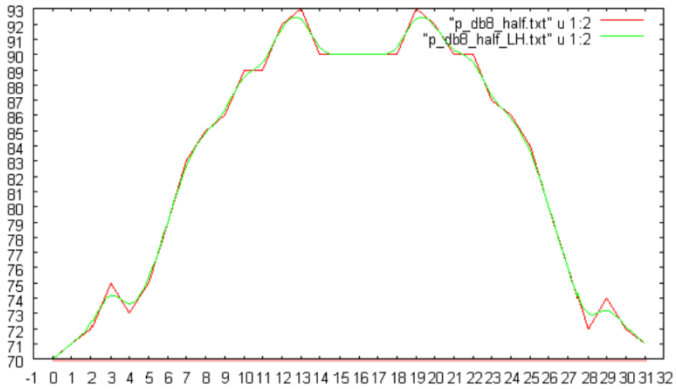


(b)Support length=4

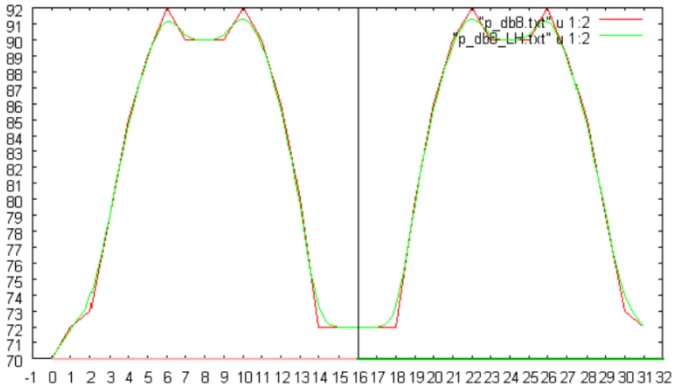


(c)Support length=8

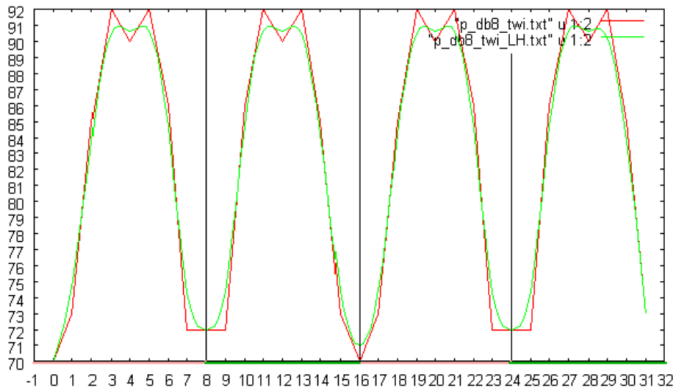
Figure 8 The number of changed pixels for the support length of 2, 4, and 8



(a)180 degree



(b)360 degree

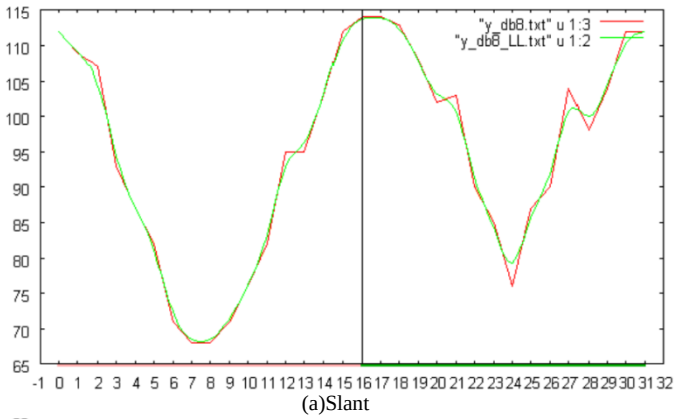


(c)720 degree

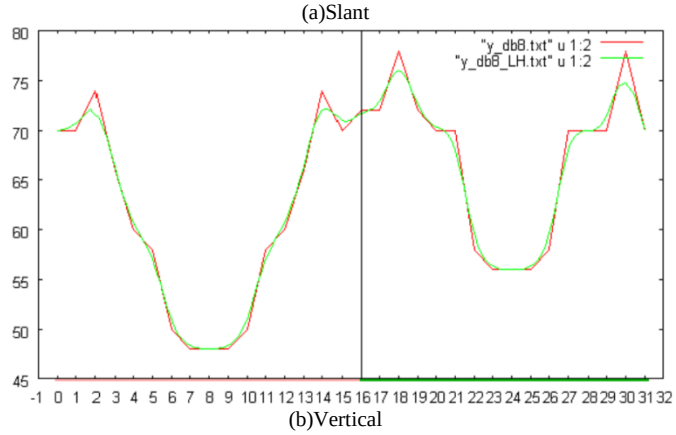
Figure 9 The number of changed pixels when the shuttlecock is rotated by 180, 360, and 720 degrees

Figure 10 (a), (b), and (c) shows the number of changed pixels in the slant, vertical, and horizontal directions, respectively, which are extracted based on the wavelet transformation with Daubechies base function of the support length of 8. In this case, the shuttlecock is rotated in pitch direction by 360 degrees. Meanwhile, Figure 11 (a), (b), and (c) shows the number of changed pixels in the horizontal direction which are extracted based on the wavelet transformation with Daubechies base function of the support length of 2, 4, and 8, respectively. In this case, the shuttlecock is rotated in pitch direction by 360 degrees.

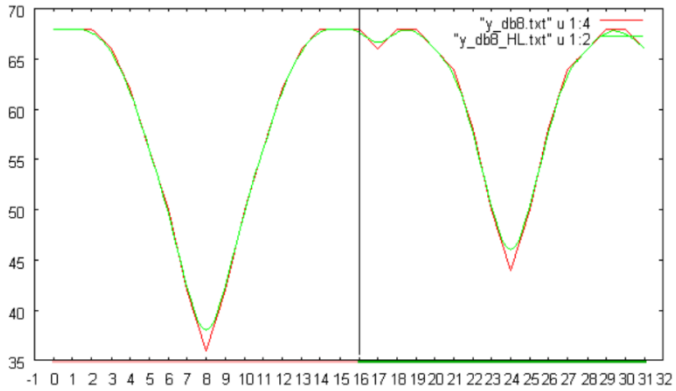
Figure 12 (a), (b), and (c) shows the number of changed pixels in horizontal direction, when the shuttlecock is rotated by 180, 360, and 720 degrees, respectively.



(a)Slant

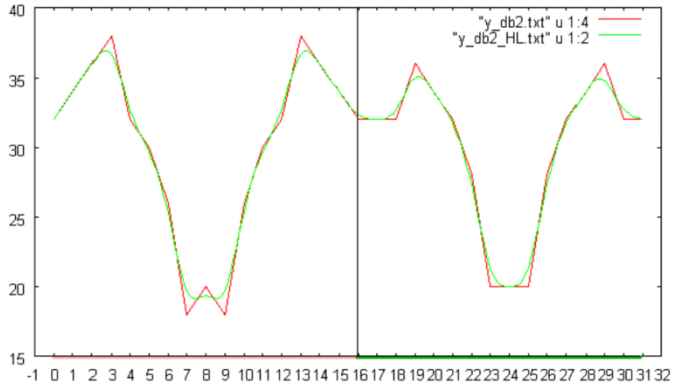


(b)Vertical

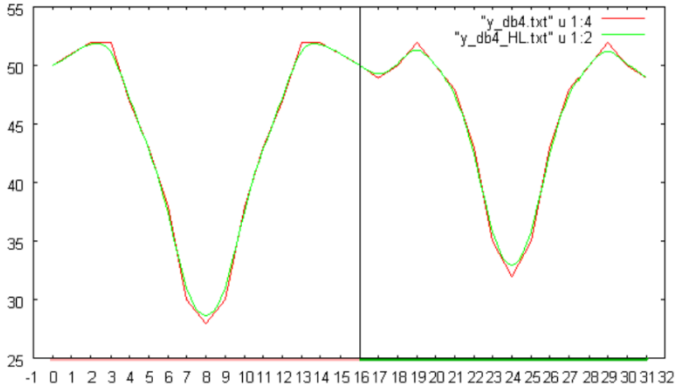


(c)Horizontal

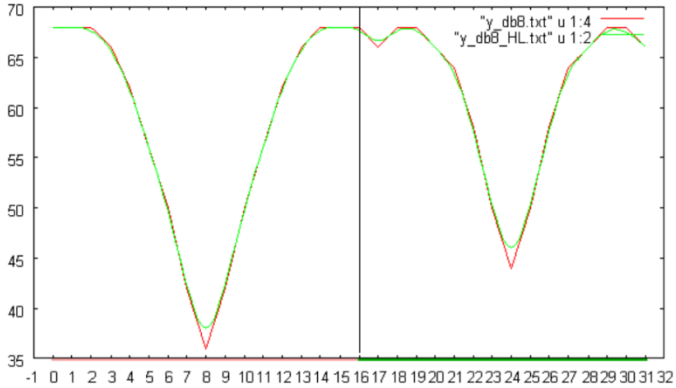
Figure 10 The number of changed pixels in the slant, vertical, and horizontal directions, respectively, which are extracted based on the wavelet transformation with Daubechies base function of the support length of 8. In this case, the shuttlecock is rotated in pitch direction by 360 degrees.



(a)Support length=2

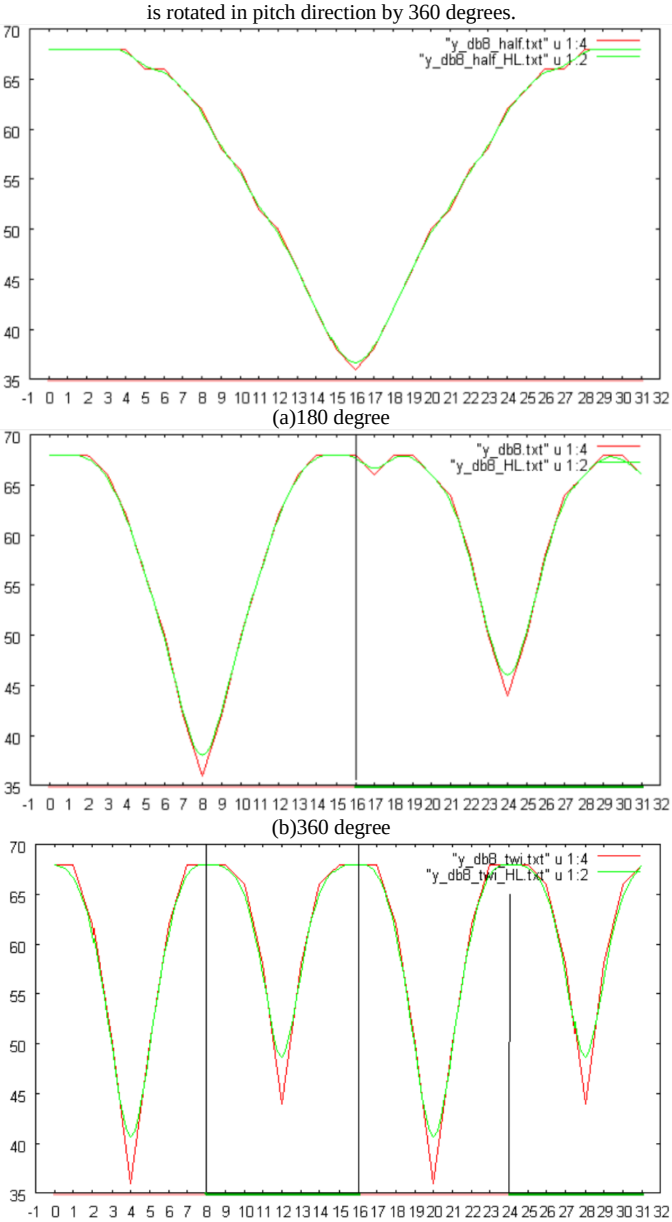


(b)Support length=4



(c)Support length=8

Figure 11 The number of changed pixels in the horizontal direction which are extracted based on the wavelet transformation with Daubechies base function of the support length of 2, 4, and 8, respectively. In this case, the shuttlecock is rotated in pitch direction by 360 degrees.



(a)180 degree

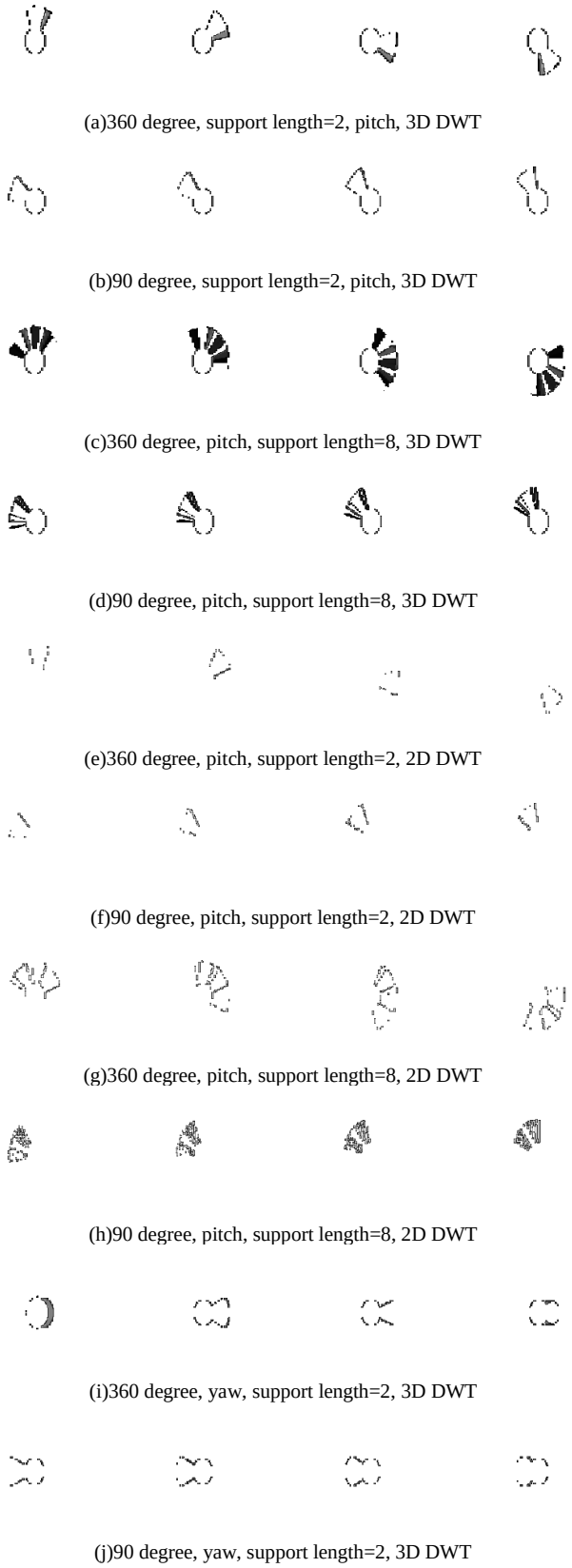
(b)360 degree

(c)720 degree

Figure 12 The number of changed pixels in horizontal direction, when the shuttlecock is rotated by 180, 360, and 720 degrees, respectively.

Figure 13 shows the edge enhanced images by 2D and 3D Discrete Wavelet Transformation: DWT with Daunechies base function of support length of 2 and 8 when the shuttlecock rotated in pitch and yaw angles with 90 and 360 degrees.

In general, the number of edge, or changed pixels extracted by 2D DWT is less than that by 3D DWT. Therefore, 3D DWT is much effective than 2D DWT. Also the number of edge, or changed pixels by DWT with support length of 2 is less than that with support length of 8. Therefore, DWT with support length of 2 is effective to extract edge, or changed pixels for the object which moves fast while DWT with support length of 8 is effective to extract edge, or changed pixels for the object which moves slowly. The terminology of fast and slowly implies the moving speed of the object referencing to the frame rate.



(a)360 degree, support length=2, pitch, 3D DWT

(b)90 degree, support length=2, pitch, 3D DWT

(c)360 degree, pitch, support length=8, 3D DWT

(d)90 degree, pitch, support length=8, 3D DWT

(e)360 degree, pitch, support length=2, 2D DWT

(f)90 degree, pitch, support length=2, 2D DWT

(g)360 degree, pitch, support length=8, 2D DWT

(h)90 degree, pitch, support length=8, 2D DWT

(i)360 degree, yaw, support length=2, 3D DWT

(j)90 degree, yaw, support length=2, 3D DWT

(k)360 degree, yaw, support length=8, 3D DWT

(l)90 degree, yaw, support length=8, 3D DWT

(m)360 degree, yaw, support length=2, 2D DWT

(n)90 degree, yaw, support length=2, 2D DWT

(o)360 degree, yaw, support length=8, 2D DWT

(p)90 degree, yaw, support length=8, 2D DWT

Figure 13 Edge enhanced images by 2D and 3D Discrete Wavelet Transformation: DWT with Daunechies base function of support length of 2 and 8 when the shuttlecock rotated in pitch and yaw angles with 90 and 360 degrees.

IV. CONCLUSION

Method for object motion characteristic estimation based on wavelet Multi-Resolution Analysis: MRA is proposed. With moving pictures, the motion characteristics, direction of translation, roll/pitch/yaw rotations can be estimated by MRA with an appropriate support length of the base function of wavelet. Through simulation study, method for determination of the appropriate support length of Daubechies base function is clarified. Also it is found that the proposed method for object motion characteristics estimation is validated.

The number of edge, or changed pixels extracted by 2D DWT is less than that by 3D DWT. Therefore, 3D DWT is much effective than 2D DWT. Also the number of edge, or changed pixels by DWT with support length of 2 is less than that with support length of 8. Therefore, DWT with support length of 2 is effective to extract edge, or changed pixels for the object which moves fast while DWT with support length of 8 is effective to extract edge, or changed pixels for the object which moves slowly. The terminology of fast and slowly implies the moving speed of the object referencing to the frame rate.

ACKNOWLEDGMENT

The author would like to thank Mr. Yuji Yamada for his effort to conduct the simulation study.

REFERENCES

[1] K. Arai K. et al., Takagi and Shimoda edt., Image Analysis Handbook, Tokyo Daigaku Shuppan-kai publishing (1991).

[2] K. Arai K., Fundamental theory for image processing, Gakujutsu-Tosho Shuppan Publishing Co., Ltd (1996).

[3] K. Arai K., Methods for Image Processing and Analysis of Earth Observation Satellite Imagery Data, Morikita Shuppan Publishing Co., Ltd (1998).

[4] K. Arai and L. Jameson, Earth observation satellite data analysis based on wavelet analysis, Morikita-Shuppan Publishing Co., Ltd (2001).

[5] K. Arai K., Java based Earth observation satellite imagery data processing and analysis, Morikita-Shuppan Publishing Co., Ltd (2002).

[6] K. Arai, K.Seto, Dynamic characteristic idetification of moving objects based on wavelet analysis, Journal of Visualization Society of Japan, 24, Suppl.1, 227-230 (2004)

[7] K.Arai, K.Seto, T. Nishikawa, Methid for dyanamic characteristic identification of rotated accelerated object based on wavelet analysis, Journal of Visualization Society of Japan, 26, Suppl.1, 145-148, 2006

[8] K.Arai, T. Nishikawa, Dangerous car identification based on wavelet analysis, Journal of Visualization Society of Japan, 27、Suppl.1、217—220、2007

[9] K.Arai, T.Nishikawa, Identification method for moving object in remote sensing satellite images based on wavelet analysis, Journal of Visualization Society of Japan, 29, Suppl.1, 163-166, 2009.

[10] K.Arai, T.Nishikawa, Method for optimum support length determinations of Daubechies base function of wavelet in moving object detection of remote sensing satellite images, Journal of Visualization Society of Japan, Journal of Japan Society of Photogrammetry and Remote Sensing, 48、3、171-179、2009.

[11] K.Arai, Yuji Yamada, Method for determination of optimum supoort length of base function of wavelet for change detection, Journal of Visualization Society of Japan, 29、Suppl.2,83-86, 2009

[12] K.Arai, Yuji Yamada, Analysis method for 3D object moements based on wavelet analysis. Journal of Visualization Society of Japan, 30,Suppl.2,287-288,2010

AUTHORS PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science, and Technology of the University of Tokyo from 1974 to 1978 also was with National Space Development Agency of Japan (current JAXA) from 1979 to 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He was appointed professor at Department of Information Science, Saga University in 1990. He was appointed councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was also appointed councilor of Saga University from 2002 and 2003 followed by an executive councilor of the Remote Sensing Society of Japan for 2003 to 2005. He is an adjunct professor of University of Arizona, USA since 1998. He also was appointed vice chairman of the Commission “A” of ICSU/COSPAR in 2008. He wrote 30 books and published 332 journal papers.

An interactive Tool for Writer Identification based on Offline Text Dependent Approach

Saranya K
Research Scholar
PSGR Krishnammal College for Women
Coimbatore, India

Vijaya MS
Associate Professor
GR Govindarajulu School Of Applied Computer Technology
Coimbatore, India

***Abstract*— Writer identification is the process of identifying the writer of the document based on their handwriting. The growth of computational engineering, artificial intelligence and pattern recognition fields owes greatly to one of the highly challenged problem of handwriting identification. This paper proposes the computational intelligence technique to develop discriminative model for writer identification based on handwritten documents. Scanned images of handwritten documents are segmented into words and these words are further segmented into characters for word level and character level writer identification. A set of features are extracted from the segmented words and characters. Feature vectors are trained using support vector machine and obtained 94.27% accuracy for word level, 90.10% for character level. An interactive tool has been developed based on the word level writer identification model.**

Keywords- Feature Extraction; Support Vector Machine; Training, Writer Identification.

I. INTRODUCTION

The significance and scope of writer identification is becoming more prominent in these days. Identification of a writer is highly essential in areas like forensic expert decision-making systems, biometric authentication in information and network security, digital rights administration, document analysis systems and also as a strong tool for physiological identification purposes.

In forensic science writer identification is used to authenticate documents such as records, diaries, wills, signatures and also in criminal justice. The digital rights administration system is used to protect the copyrights of electronic media. Two broad categories of biometric modalities are: physiological biometrics that perform person identification based on measuring a physical property of the human body (e.g. fingerprint, face, iris, retinal, hand geometry) and behavioral biometrics that use individual traits of a person's behavior for identification (e.g. voice, gait, signature, handwriting). Hence writer identification falls under the category of behavioral biometrics. Handwritten document analysis is applied in fields of information retrieval either textually or graphically [1].

Writer identification mode can be generally classified into two types as online and offline. In online, the writing behavior is directly captured from the writer and converted to a sequence of signals using a transducer device but in offline the handwritten text is used for identification in the form of scanned images. Off-line writer identification is extensively considered as more challenging than on-line because it contains more information about the writing style of a person, such as pressure, speed, angle which is not available in the off-line mode.

Writer identification approaches can be categorized into two types: text-dependent and text-independent methods. In text-dependent methods, a writer has to write the identical text to perform identification but in text independent methods any text may be used to establish the identity of writer [2].

Various approaches and techniques have been proposed so far for writer identification. Writer identification using connected component contours codebook and its probability density function was proposed in [3]. This paper exhibits better identification rates by combining connected-component contours with an independent edge-based orientation and curvature PDF. In [4], eleven macro-features and micro-features have been used for writer identification. Authors in [8] have used a set of features extracted from lines of text correspond to visible characteristics of the writing such as width, slant, height of the three main writing zones and also features based on the fractal behavior of the writing for writer identification. A system for writer identification using textural features derived from the gray-level co-occurrence matrix and Gabor filters has been described in [12]. In the research work [14], Morphological features obtained from transforming the projection of the thinned writing have been computed and used for writer identification. A HMM based approach for writer identification and verification built an individual recognizer for each writer and train it with text lines of writer was proposed in [15]. A system developed in [16] for writer identification and verification takes two pages of handwritten text as input and determines whether the same writer has written those two pages and features like character height, stroke width, writing slant and skew, frequency of loops and blobs have been used.

This research proposes text dependent writer identification based on scanned images of English handwriting. The scanned images are segmented into words and these words are further segmented into characters on which pre-processing and features extraction tasks are performed.

Features like edge based features, word measurements, moment invariants used in the existing research work are taken

into account. Edge based features are computed using edge detected image. Edge based directional distribution and edge hinge distributions are two edge based features. Features such as length of the word, height of the word, height from baseline to upper edge, height from baseline to lower edge, ascender and descender baseline are word measurement features. Moment invariant calculates a set of seven moments for a given image.

This paper includes additional features that were not taken into consideration in [1]. They are the character level features like aspect ratio, loops, junctions and end points. The proposed work is implemented using Support Vector Machine, a supervised learning technique. Additionally an interactive tool has been developed when compared to existing works.

II. Proposed Writer Identification Model

The basic property of handwriting is that there exists writer invariant which makes writer identification possible. The writer's invariants reflecting the writing style or writing individuality of handwriting can be defined as the set of similar patterns. Also, two samples of a writer cannot be same. Hence accurate prediction of writer is highly important and challenging task. Hence it is proposed to design and develop a tool for recognizing a writer based on his / her handwriting using pattern recognition technique.

The essential tasks of writer identification are data acquistion, scanning, segmentation, feature extraction, training and writer recognition. The architecture of the proposed system is shown in Fig.1.

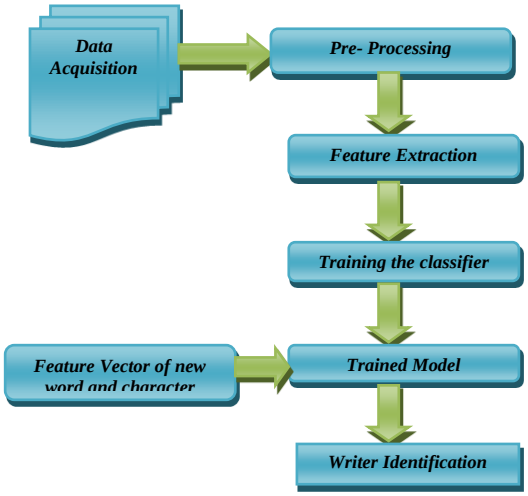


Fig.1 System Architecture

A. Data Aquistion

The data acquistion is an important task in writer identification. In order to acquire an acceptable data, identical words written using the same pen by different writers have been used. Words that have been collected are not case sensitive and the words are scanned using scanner of resolution 300 dpi. A total of 1000 JPEG text images from 10 writers of different age groups and 100 words per writer are obtained.

B. Pre-Processing

In pre-processing segmentation, noise removal, binarization, edge detection and thinning operations are performed.

Segmentation: The scanned handwritten document containing 100 words of a writer is segmented into words using the edge pixels. These words are further segmented into characters till reaches 100 characters per writer.

Noise Removal: Noise in an image is removed using median filtering.

Binarization: This converts gray scale image into binary image using Otu's method.

Edge Detection: Edges in the binary image are detected using sobel method.

Thinning: Morphological operations are used for thinning the binary image.

C. Feature Extraction

Feature extraction plays a vital role in improving the classification effectiveness and computational efficiency. A set of distinctive features describing the writing style and writer's invariance is extracted to form a feature vector. The features are described below.

1) Edge Direction Distribution

In edge direction distribution, first the edge of the binary image is detected using Sobel detection method. The edge detected image are labelled using 8- connected pixel neighbouhood. Then the number of rows and columns in an binary image is found using size function. Next, the first black pixel in an image is found and this pixel is considered as center pixel of the square neighbourhood. Then the black edge is checked using logical AND operator in all direction starting from the center pixel and ending in any one of the edge in the square. In order to avoid redundancy the upper two quadrants in the neighbourhood is checked because without on-line information, it is difficult to identify the way the writer travelled along the edge fragment. This will gives us "n" possible angles. Subsequently, the verified angles of each pixel are counted into n-bin histogram which is then normalized to a probability distribution which in turn gives the probability of an edge fragment oriented in the image at the angle measured from the horizontal. Here "n" is taken as 4, 8, 12, and 16.

2) Edge Hinge Distribution

To capture the curvature of ink trace, which is very distinctive for different writer, edge hinge distribution is needed, which is calculated with the help of local angles along the edges. Edge hinge feature considers two edge fragments emerging from center pixel and, subsequently, joint probability distribution of the orientations of the two fragments of a 'hinge' are calculated. Finally, normalized histogram gives the joint probability distribution for "hinged" edge fragments oriented at the angles 1 and 2. The orientation is counted in 16 directions for a single angle. From the total number of combinations of two angles only non- redundant values are considered and the common ending pixels are eliminated.

3) Run Length Distribution

Run lengths are determined on binarized image taking into consideration either the black pixels consistent to the ink trace or the white pixels matching to the background. Scanning procedures are of two types: horizontal along the rows of the image and vertical along the column of the image. Next, the probability distribution is interpreted by using the normalized histogram of run lengths. Orthogonal information to the directional features is obtained by using the run lengths.

4) Auto-Correlation

Auto-correlation function identifies the presence of predictability in writing. By giving the offset value, every row of the image is shifted onto itself. Then the normalized dot product is found between the original row and the shifted row. Auto-correlation function is computed for all rows and the sum is normalized to obtain a zero-lag correlation of 1.

5) Entropy

Entropy provides the average information of an image such as luminance, contrast and pixel value. It is calculated using the formula:

$$E = H[p(g)] - \sum_{j-1}^{j} p(j)H[p_j(g)]$$

6) Moment Invariants

Geometric moment invariant is commonly used in pattern recognition. A distinctive set of features calculated for an object must be able to identify the same object with another possible different size and orientation. Moment invariants can be used to recognize object when the object is changed in transformations. Here the following seven moments are computed.

$$M1 = \eta 20 + \eta 02,$$

$$M2 = (\eta 20 - \eta 02)^2 + (2\eta 11)^2,$$

$$M3 = (\eta 30 - 3\eta 12)^2 + (3\eta 21 - \eta 03)^2,$$

$$M4 = (\eta 30 + \eta 12)^2 + (\eta 21 + \eta 03)^2,$$

$$M5 = (\eta 30 - 3\eta 12)(\eta 30 + \eta 12)\left[(\eta 30 + \eta 12)^2 - 3(\eta 21 + \eta 03)^2\right] + (3\eta 21 - \eta 03)(\eta 21 + \eta 03)\left[3(\eta 30 + \eta 12)^2 - (\eta 21 + \eta 03)^2\right],$$

$$M6 = (\eta 20 - \eta 02)\left[(\eta 30 + \eta 12)^2 - (\eta 21 + \eta 03)^2\right] + 4\eta 11(\eta 30 + \eta 12)(\eta 21 + \eta 03),$$

$$M7 = (3\eta 21 - \eta 03)(\eta 30 + \eta 12)\left[(\eta 30 + \eta 12)^2 - 3(\eta 21 + \eta 03)^2\right] - (\eta 30 + 3\eta 12)(\eta 21 + \eta 03)\left[3(\eta 30 + \eta 12)^2 - (\eta 21 + \eta 03)^2\right].$$

7) Length

Length of the word and character is found by successively penetrating each column in the binary image to find the first and last pixels in the image and store their column numbers. The length of the image is calculated by subtracting the column number of last pixel to the column number of first pixel.

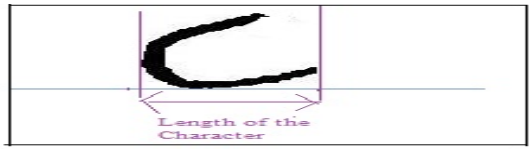


Fig. 2 Length of the character

8) Height

Height of the word and character is found by consecutively probing each row in the binary image. The first and last pixels of the image are found and the corresponding row numbers are stored. The height of the image is computed by subracting from the row number of last pixel to the row number of first pixel.

Fig. 3 Height of the character

9) Area

Area of the word and character is calculated as the product of height and length.

10) Height from baseline to upper edge

The height of the text from baseline to upperedge is calculated by first determining the baseline position of the image. This is functioned by casting an array where the index is row number in the image.

Then, the number of black pixels in each row is calculated and the results are stored in an array. After completing the entire image, the maximum value of the array is identified and the corresponding row number is stored as the baseline. The length of the image from the baseline to the upper edge is computed by subtracting the row number of first pixel in the image from its row number of the baseline.

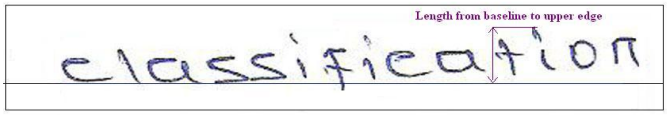


Fig. 4 Length from baseline to upper edge

11) Height from baseline to lower edge

The height of the binary image from the baseline to the lower edge is determined by calculating the baseline row number, as above.

Then, the row number of the last pixel of the image is considered. The height of the image from the baseline to the lower edge is calculated by subtracting the last pixel row number to the baseline row number.

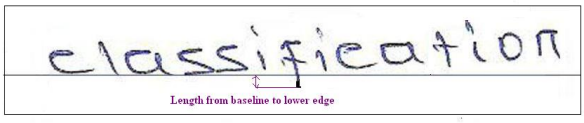


Fig. 5 Length from baseline to lower edge

12) Ascender and descender baseline

Ascender baseline is the first non-zero value of column and the descender baseline is the last non-zero value of column of the vertical histogram of the line.

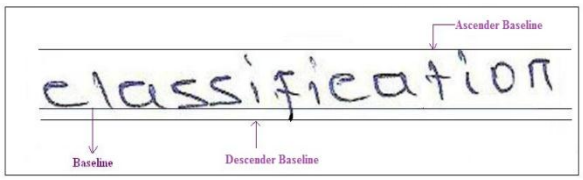


Fig. 6 Ascender and Descender Baseline

13) Aspect Ratio

Aspect ratio is considered as one of the global features in writer identification. It is calculated as ratio of width to height.

14) End Points

End-points contain only one pixel in their 8-pixel neighborhood. It is computed using end point function which gives the number of end points in the thinned image.

15) Junctions

Junctions occur where two strokes meet or cross and are found in the skeleton as points with more than two neighbors. It produces number of junctions, positions of each junction, angle and distance between the junctions of the thinned image.

16) Loops

The loops of a character are the major distinguishing feature for many writers. The loop function gives loop length, angle of loop, position of the loop, area and average radius of the loop of the edge image.

17) Slope Angle

The slope of angle **A** is the ratio of height to length. In geometry, it is also referred to as the tangent of the angle **A** and denoted by tan(**A**), which gives us the slope angle.

18) Slant Angle

It is the angle of the word forms against the baseline. It is estimated on structural features by maxima and minima of the word are detected and targets uniform slant angle estimation.

Thus a total of 26 features are extracted from a single word image. The same set of 26 features is extracted from each character image. Finally two independent training dataset each consisting of 1000 instances has been developed using MATLAB.

D. Support Vector Machine

Support vector machine is a training algorithm for learning classification and regression rules from data. SVM is very suitable for working accurately and efficiently with high dimensionality feature spaces. The machine is presented with a set of training examples, (xi,yi) where the xi are the real world data instances and the yi are the labels indicating which class the instance belongs to. For the two class pattern recognition problem, yi = +1 or yi = -1.

SVMs construct a hyperplane that separates two classes and tries to achieve maximum separation between the classes. Separating the classes with a large margin minimizes a bound on the expected generalization error. The simplest model of SVM called Maximal Margin classifier, constructs a linear separator (an optimal hyperplane) given by w T x - γ = 0 between two classes of examples. The free parameters are a vector of weights **w,** which is orthogonal to the hyperplane and a threshold value γ. These parameters are obtained by solving the following optimization problem using Lagrangian duality.

$$\text{Minimize} \quad = \frac{1}{2}\|W\|^2$$

$$\text{subject to} \quad D_{ii}\left(W^T X_i - \gamma\right) \geq 1, i = 1,.....,l.$$

Where D_{ii} corresponds to class labels +1 and –1. The instances with non-null weights are called support vectors. In the presence of outliers and wrongly classified training examples it may be useful to allow some training errors in order to avoid over fitting. A vector of slack variables ξi that measure the amount of violation of the constraints is introduced and the optimization problem referred to as soft margin is given below.

$$\underset{W,\gamma}{Minimize} = c\sum_{i=1}^{l}\varepsilon_i + \frac{1}{2}\|W\|^2$$

$$\text{subject to} \quad \underset{\varepsilon_i \geq 0}{D_{ii}}\left(W^T X_i - \gamma\right) + y_i \geq 1, i = 1,.....,l.$$

In this formulation the contribution to the objective function of margin maximization and training errors can be balanced through the use of regularization parameter c. The following decision rule is used to correctly predict the class of new instance with a minimum error.

$$f(X) = \text{sgn}\left[W^T x - \gamma\right]$$

The advantage of the dual formulation is that it permits an efficient learning of non–linear SVM separators, by introducing kernel functions. Technically, a kernel function calculates a dot product between two vectors that have been (non linearly) mapped into a high dimensional feature space. Since there is no need to perform this mapping explicitly, the training is still feasible although the dimension of the real feature space can be very high or even infinite. The parameters are obtained by solving the following non-linear SVM formulation (in Matrix form),

$$\text{Minimize } L_D(\mathbf{u}) = \frac{1}{2}\mathbf{u}^T\mathbf{Q}\mathbf{u} - \mathbf{e}^T\mathbf{u}$$

$$\mathbf{d}^T\mathbf{u} = 0 \;,\; \mathbf{0} \leq \boldsymbol{u} \leq \boldsymbol{Ce}$$

Where $\boldsymbol{Q = DKD}$ and $\boldsymbol{K}$ - the Kernel Matrix. The kernel function K (AAT) (polynomial or Gaussian) is used to construct hyperplane in the feature space, which separates two classes linearly, by performing computations in the input space. The decision function is given by

$$f(X) = \text{sgn}\left(K\left(x, x_i^T\right) * u - \gamma\right)$$

where, u - the Lagrangian multipliers.

When the number of class labels is more than two, the binary SVM can be extended to multi class SVM. The indirect methods for multiclass SVM is one versus rest method. For each class a binary SVM classifier is constructed, discriminating the data points of that class against the rest.

Thus in case of N classes, N binary SVM classifiers are built. During testing, each classifier yields a decision value for the test data point and the classifier with the highest positive decision value assigns its label to the data point. The comparison between the decision values produced by different SVMs is still valid because the training parameters and the dataset remain the same.

III. Experiment And Results

Two independent experiments have been carried out, one for word level writer identification and another for character level writer identification. Two training datasets that have been developed for word level and character level are used for implementation. The datasets are normalized using min-max normalization and the normalized datasets is used for learning SVM.

The normalized datasets are trained independently using SVM^{light} for support vector machine with linear, polynomial and RBF kernels with different parameter value for C, where C is the regularization parameter. The parameters d and gamma are associated with polynomial kernel and RBF kernel respectively. The performance of trained models is evaluated using 10-fold cross validation for its predictive accuracy and the learning time. The prediction accuracy is the ratio of number of correctly classified instances in the test dataset and the total number of test cases.

A. Word Level Writer Identification

The regularization parameter C is assigned values between 0.5 and 50 for linear kernel. For polynomial and RBF kernels the value for C is assigned as 0.5, 1 and 5, d is assigned from 1 to 4 and g is taken from 0.5 to 5 respectively. It is found that the model performs better for the value C = 5**.**

The results of word level writer identification model based on SVM with linear kernel are shown in Table I.

TABLE I. RESULTS OF LINEAR SVM FOR WORD LEVEL

C	Prediction Accuracy (%)	Time Taken (in Secs)
0.5	62.50	0.01
1	65.10	0.01
5	70.83	0.02
10	70.31	0.02
15	75.00	0.01
20	70.65	0.03
25	69.27	0.01
30	68.75	0.04
35	73.95	0.02
40	71.35	0.03
45	77.60	0.01
50	73.43	0.02

The results of word level writer identification model based on SVM with polynomial kernel are shown in Table II.

TABLE II. RESULTS OF SVM WITH POLYNOMIAL KERNEL FOR WORD LEVEL

C	d	Prediction Accuracy (%)	Time Taken (in Secs)
0.5	**1**	78.12	1.18
	2	93.75	0.77
	3	91.66	2.90
	4	93.85	0.76
1	**1**	66.66	1.58
	2	72.91	2.35
	3	94.07	2.26
	4	93.22	0.96
5	**1**	71.87	1.79
	2	78.64	2.02
	3	94.27	3.49
	4	91.69	1.52

The results of word level writer identification model based on SVM with RBF kernel are shown in Table III.

TABLE III. RESULTS OF SVM WITH RBF KERNEL FOR WORD LEVEL

C	g	Prediction Accuracy (%)	Time Taken (in Secs)
0.5	**0.5**	84.37	0.04
	1	89.58	0.03
	1.5	91.10	0.04
	2	92.18	0.04
	2.5	91.19	0.05
	3	91.05	0.04
1	**0.5**	84.65	0.06
	1	89.15	0.08
	1.5	91.14	0.06
	2	92.08	0.07
	2.5	91.66	0.06
	3	91.86	0.06
5	**0.5**	78.12	0.14
	1	89.40	0.12
	1.5	91.14	0.06
	2	92.28	0.08
	2.5	91.66	0.07
	3	91.69	0.06

B. Character Level Writer Identification

The training dataset that has been developed using character level images is used here for SVM learning. The

parameter settings for character level training are same as word level training.

The results of character level writer identification model based on SVM with linear kernel are shown in Table IV.

TABLE IV. RESULTS OF LINEAR SVM FOR CHARACTER LEVEL

C	Prediction Accuracy (%)	Time Taken (in Sec)
0.5	62.08	0.01
1	60.18	0.02
5	66.35	0.02
10	72.03	0.01
15	63.98	0.03
20	64.45	0.03
25	70.61	0.01
30	72.58	0.04
35	61.13	0.03
40	74.40	0.02
45	67.77	0.05
50	72.98	0.03

The results of character level writer identification model based on SVM with polynomial kernel are shown in Table V.

TABLE V. RESULTS OF SVM WITH POLYNOMIAL KERNEL FOR CHARACTER LEVEL

C	d	Prediction Accuracy (%)	Time Taken (in Sec)
0.5	**1**	64.58	1.25
	2	68.75	1.95
	3	66.67	2.06
	4	88.54	3.56
1	**1**	73.43	1.45
	2	71.35	2.09
	3	77.60	3.12
	4	79.16	4.45
5	**1**	61.45	1.57
	2	63.02	2.20
	3	83.85	5.65
	4	90.10	8.03

The results of character level writer identification model based on SVM with RBF kernel are shown in Table VI.

TABLE VI. RESULTS OF SVM WITH RBF KERNEL FOR CHARACTER LEVEL

C	G	Prediction Accuracy (%)	Time Taken (in Sec)
0.5	**0.5**	65.10	0.05
	1	76.56	0.06
	1.5	81.52	0.06
	2	82.07	0.04
	2.5	83.03	0.06
	3	84.21	0.05
1	**0.5**	52.60	0.13
	1	76.56	0.07
	1.5	81.77	0.06
	2	82.60	0.06
	2.5	83.00	0.07
	3	84.37	0.06
5	**0.5**	52.60	0.12
	1	71.35	0.12
	1.5	79.16	0.12
	2	85.41	0.13
	2.5	82.81	0.06
	3	83.33	0.06

C. Comparative Analysis

The average and comparative performance of SVM's with various kernels for word level and character level writer identification is given in the Table. VII and shown in Fig.7 and Fig.8.

TABLE VII. AVERAGE PERFORMANCE OF THREE MODELS

SVM Kernels	Word Level Writer Identification			Character Level Writer Identification		
	Para-meter	**Accuracy (%)**	**Time Taken (in secs)**	**Para-meter**	**Accuracy (%)**	**Time Taken (in secs)**
Linear	C = 45	77.60	0.01	C = 40	74.40	0.02
RBF	C = 5, g = 2	92.28	0.08	C = 5, g = 2	85.41	0.13
Polynomial	C = 5, d = 3	94.27	3.49	C = 5, d = 4	90.10	8.03

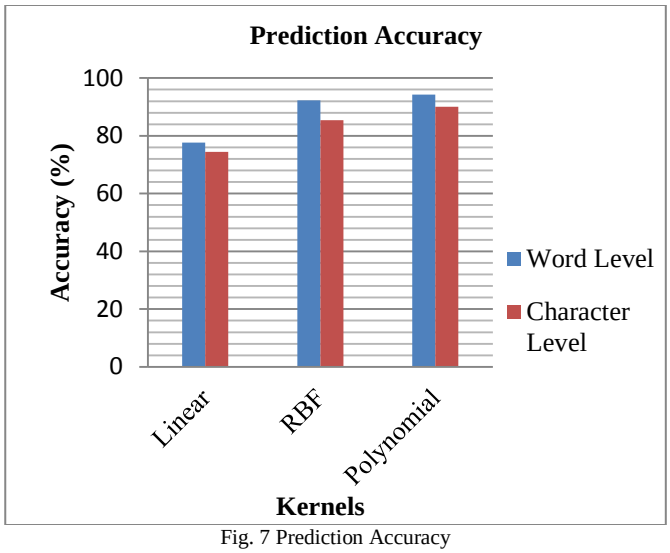


Fig. 7 Prediction Accuracy

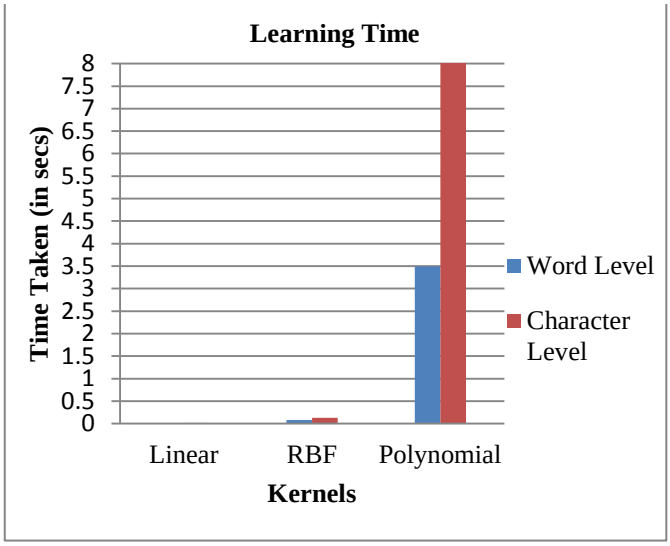


Fig. 8 Learning Time

From the above comparative analysis the predictive accuracy shown by SVM with polynomial kernel is higher than the linear and RBF kernel. The time taken to build the model using SVM with polynomial kernel is more, than linear and RBF kernel. As far as the writer identification is concerned accuracy plays major role than learning time in identifying the writer.

Also it is found that about 94.27% predictive accuracy for word level writer identification and 90.10% predictive accuracy for character level writer identification are shown by SVM based prediction model. Hence it is found that word level writer identification performs better than character level writer identification. The SVM polynomial based model which is producing higher accuracy is taken into consideration for developing writer identification tool.

IV. Writer Identification Tool

An interactive writer identification tool is developed using MATLAB by incorporating the SVM model with GUI. This tool is used to predict the individual when his/her offline handwriting of a new word is given as input. Screenshots for writer identification tool are shown in Fig. 9 to 12.

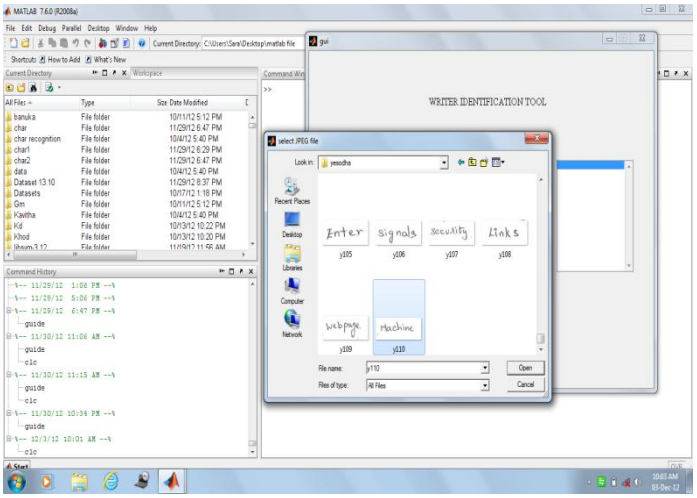


Fig. 9 Choosing the word

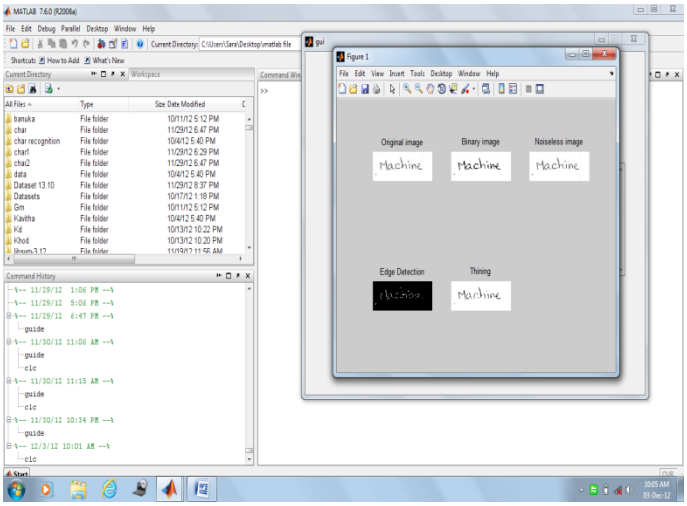


Fig. 10 After Pre-Processing

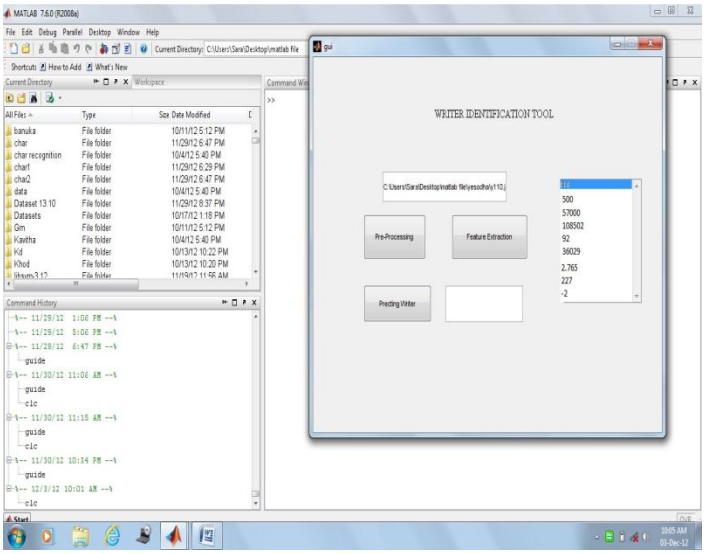


Fig. 11 Feature Extraction

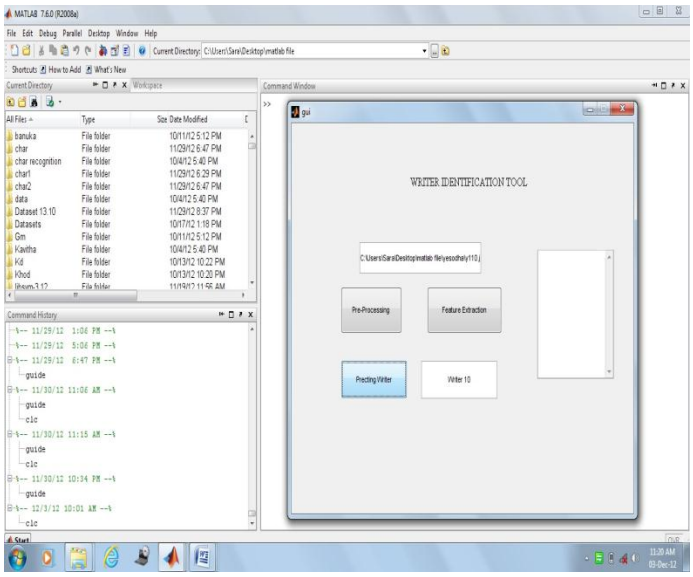


Fig. 12 Predicting the writer

V. CONCLUSION

This paper describes the modeling of writer identification problem as classification task. Two independent dataset has been prepared in order to facilitate training and implementation. The outcome of the experiments indicates that the SVM with polynomial kernel for word level writer identification predicts the writer of the handwritten document more accurately than the other models. Based on this model writer identification tool has been developed to predict the writer. Online text independent approach can be implemented as a future work,

REFERENCES

[1] Al-Maʿadeed S, Mohammed E, AlKassis D, Al-Muslih F, "Writer identification using edge-based directional probability distribution features for Arabic words," IEEE/ACS International Conference on Computer Systems and Applications, pp.582-590, 2008.

[2] Al-Dmour A, Zitar RA, "Arabic writer identification based on hybrid spectral-statistical measures," Journal of Experimental and Theoretical Artificial Intelligence, vol. 19, no. 4, pp. 307–332, 2007.

[3] Andreas Schlapbach and Horst Bunke, "*Off-lineWriter Identification Using Gaussian Mixture Models"*, 18th International Conference on Pattern Recognition, 2006

[4] Bulacu M and Schomaker L, "Text-independent writer identification and verification using textural and allographic features," IEEE Trans. on Pattern Analysis and Machine Intelligence, pp.701–717, April 2007.

[5] Cha SH and Srihari S, "Writer identification: statistical analysis and dichotomizer," Springer LNCS 1876, pp. 123-132, 2000.

[6] Crammer K and Singer Y, *"On the algorithmic implementation of Multiclass SVMs"*, JMLR, 2001

[7] Cristianini N and Shawe-Taylor J, "An Introduction to SupportVector Machines," Cambridge University Press, 2000.

[8] Helli B and Moghaddam ME, "A text-independent Persian Writer Identification based on Feature Relation Graph", Pattern Recognition, vol. 43, pp. 2199–2209, 2010.

[9] Mallikarjunaswamy BP, Karunakara K, **"**Writer Identification based on offline Handwritten Document Images in Kannada language using Empirical Mode Decomposition method," International Journal of Computer Applications, Vol. 30, No. 6, September 2011.

[10] Marti UV, Messerli N and Bunke H, "Writer identification using text line based features," IEEE Proc. of 6th Int. conf. on Document Analysis and Recognition, pp. 101-105, 2001.

[11] Said HES, Peake GS, Tan TN and Baker KD, "Personal identification based on handwriting," Pattern Recognition, vol. 33, pp. 149-160, 2000.

[12] Schlapbach A and Bunke H, "Using HMM based recognizers for writer identification and verification," IEEE Proc. Of 9th Int. Workshop on Frontiers in Handwriting Recognition, pp. 167-172, 2004.

[13] Schomaker L and Bulacu M, "Automatic writer identification using connected component contours and edge-based features of uppercase western script", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, pp. 787–798, June 2004.

[14] Shahabi F and Rahmati M, "A New Method for Writer Identification of Handwritten Farsi Documents," 10th International Conference in Document Analysis and Recognition, pp. 426-430, 2009.

[15] Sreeraj M, Idicula SM, "Identifying Decisive Features for Distinctive Analysis of Writings in Malayalam," IMACST: vol. 2, No. 1, may 2011.

[16] Srihari S, Cha S, Arora H, and Lee S, "Individuality of Handwriting," J. Forensic Sciences, vol. 47, no. 4, pp. 1-17, 2002.

[17] Ubul K, et al., "Research on Uyghur off-line handwriting-based writer identification," 9th International Conference in Signal Processing, pp. 1656-1659, 2008.

[18] Xin Li, Xianliang Wang and Xiaoqing Ding,*"An Off-line Chinese Writer Retrieval System Based on Text-sensitive Writer Identification"*, 18th International Conference on Pattern Recognition, 2006.

[19] Yan Y, Chen Q, Deng W and Yuan F, "Chinese Handwriting Identification Based on Stable Spectral Feature of Texture Images," International Journal of Intelligent Engineering and Systems, Vol.2, No.1, 2009.

[20] Zois EN and Anastassopoulos V, "Morphological waveform coding for writer identification," Pattern Recognition, vol. 33(3), pp. 385-398, 2000.

Multi-modal Person Localization And Emergency Detection Using The Kinect

Georgios Galatas
Computer Sci. and Eng. Dept.
University of Texas at Arlington
Arlington, TX, USA
IIT, NCSR Demokritos
Athens, Greece

Shahina Ferdous
Computer Sci. and Eng. Dept.
University of Texas at Arlington
Arlington, TX, USA

Fillia Makedon
Computer Sci. and Eng. Dept.
University of Texas at Arlington
Arlington, TX, USA

***Abstract*—Person localization is of paramount importance in an ambient intelligence environment since it is the first step towards context-awareness. In this work, we present the development of a novel system for multi-modal person localization and emergency detection in an assistive ambient intelligence environment for the elderly. Our system is based on the depth sensor and microphone array of 2 Kinect devices. We use skeletal tracking conducted on the depth images and sound source localization conducted on the captured audio signal to estimate the location of a person. In conjunction with the location information, automatic speech recognition is used as a natural and intuitive means of communication in order to detect emergencies and accidents, such as falls. Our system attained high accuracy for both the localization and speech recognition tasks, verifying its effectiveness.**

Keywords-localization; multi-modal; Kinect; speech recognition; context-awareness; 3-D interaction

I. INTRODUCTION

An assistive ambient intelligence environment is a smart space that aids the inhabitants with its embedded technology. For achieving this goal, activity recognition and emergency detection, performed in a natural and unintrusive way, are of utmost importance. The most decisive step for effective activity recognition is accurate and robust person localization. By utilizing the location of the person in a domestic setting, the related activity can be derived. In addition, an elderly-oriented assistive environment must also be able to detect emergencies and accidents, such as falls, in order to issue a distress signal. This additional role increases requirements for information redundancy and reliability. Our novel system uses information from multiple sensors in order to ensure reliable localization of the inhabitant as well as emergency detection, offering speech recognition as a means of natural interaction.

Applications that rely on localization such as surveillance and monitoring of assistive daily living (ADL) commonly use video cameras as an affordable and abundant source of information. Many approaches based on either a single camera or multiple cameras have been proposed in the literature. In single camera setups, discriminative appearance affinity models [1] and level-set segmentation [2] have been used for tracking, while other approaches based on tracking-by-detection exist [3, 4]. In multi-camera setups, stereo-vision is employed in order to introduce depth perception. In [5], color histograms of the detected persons are used, while in [6] heuristic and probabilistic tracking is used to determine the location of a person in the 3-D space. Nevertheless, the segmentation and tracking problems can be very challenging, thus hindering the system's reliability in a camera-only setup. In an attempt to improve performance, multi-modal person localization has become a significant research area in recent pervasive assistive applications. Common approaches combine cameras with microphone arrays and other sensors for localization and activity detection [7], using particle filtering for data fusion. In addition to activity recognition, some systems incorporate emergency detection and most notably fall detection. Fall detection has triggered the interest of researchers, since falls account for over 75% of domestic accidents for adults over 75 [8]. Furthermore, over 30% of adults over 65 fall at least once a year [9], making them the most common cause of injury death [10] with a direct cost of $30 billion [11]. Although the vast majority of fall detection systems use solely cameras [12, 13], some systems use a combination of sensors [14].

Nevertheless, the use of cameras in all the aforementioned implementations can be considered intrusive when used domestically. Furthermore, fall detection systems do not account for other hazards that may arise in a domestic setting and may require additional modalities to increase reliability. However, our system uses the depth sensor of a Kinect device as the main source of localization information. This active sensor is able to accurately measure the position of the person in the 3-D space. At the same time, the video information is not captured, making this approach less intrusive than using video cameras. In addition, the Kinect device incorporates a microphone array capable of localizing sounds. In our system, we use 2 Kinect devices to capture the audio signal for both improving localization by introducing an additional modality as well as enabling natural interaction by means of speech recognition. Speech recognition is also used in order to detect emergencies independently of the localization module.

In the following sections we will present the architecture and operation of our system for the person localization and emergency detection tasks, the experimental setup and finally our concluding remarks.

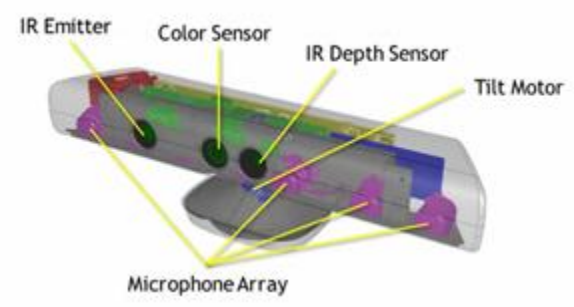


Figure 1. The Microsoft Kinect device.

II. The Kinect Device

The Microsoft Kinect (fig. 1) is a new device mainly used for gesture recognition. It is based on the PrimeSensor design [15] and it incorporates a color camera, a depth sensor and a microphone array. Depth images are acquired using the structured light technique. According to this method, a laser beam passes through a grating, and is split into different beams. The beams are then reflected from an object in the device's field of view (FOV) and captured by an infra-red sensor, making it possible to calculate the distance of the object using triangulation [16]. The microphone array is comprised of 4 microphones, enabling sound source localization. For our application, we implemented the least intrusive setup possible by capturing data only from the depth sensor and the microphone array, without capturing the actual color video data.

III. System Architecture

The architecture of our system is modular, comprising of 3 main components as shown in fig. 2. Communication between the modules is based around the Joint Architecture for Unmanned Systems (JAUS) [17], originally developed by the U.S. Department of Defense, to govern the way that unmanned systems are designed. The user datagram protocol (UDP) is used for inter-module communications, which increases the level of interoperability, allowing new software modules to be easily integrated in the system or existing modules to be installed on different systems. Input is provided by 2 Kinect devices. One of them is considered as primary, capturing both a stream of depth images and audio, while the secondary captures only audio for performing sound localization. Interfacing with the Kinect is carried out using the MS software development kit (SDK) v1.0 [18]. The 3 modules 1) skeletal tracking based localization, 2) audio localization and 3) automatic speech recognition (ASR) are described in detail in the following paragraphs.

A. Skeletal Tracking Based Localization Module

Skeletal tracking is used in our system in order to detect and track a person in the FOV of the sensor, as s/he moves in the smart space and it was implemented using the MS Kinect SDK. Initially, the moving person is detected, then her/his center of mass is determined and finally a skeletal model is fitted. The detected skeleton has a unique identifier for a specific session and is defined by the 3-D coordinates of its 20 joints $< X_{di}, Y_{di}, Z_{di} >$, expressed in meters.

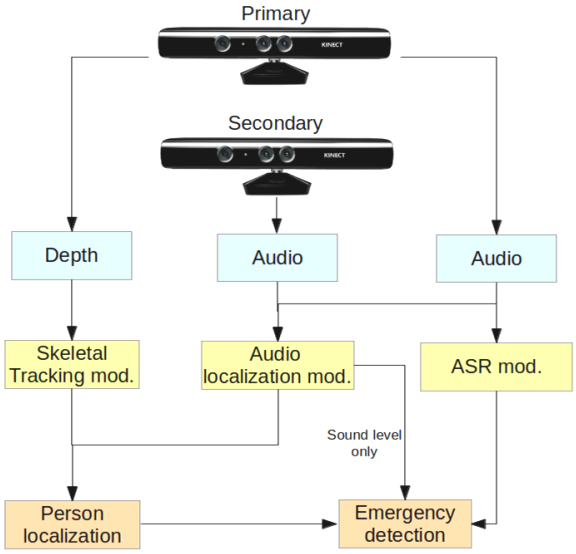


Figure 2. System architecture showing the 3 modules and 2 operation modes.

Each joint can be at any of the three associated states: 1) tracked, 2) not-tracked and 3) inferred. Furthermore, two kinds of filters are applied to the joint coordinates due to the nature of the captured data, 1) high frequency jitter and 2) temporary spikes rejection. Although the infrastructure for tracking the joints of 2 skeletons and the center of mass of 4 additional people exists, the main scope of our system is to monitor an elderly inhabitant of an assistive environment when not supervised, so at most 2 tracked skeletons are considered. Localization using such skeletal tracking is very accurate and unintrusive since we only utilize the coordinates calculated from the depth sensor feed. A visualization of the operation of the skeletal tracker is shown in fig. 3.

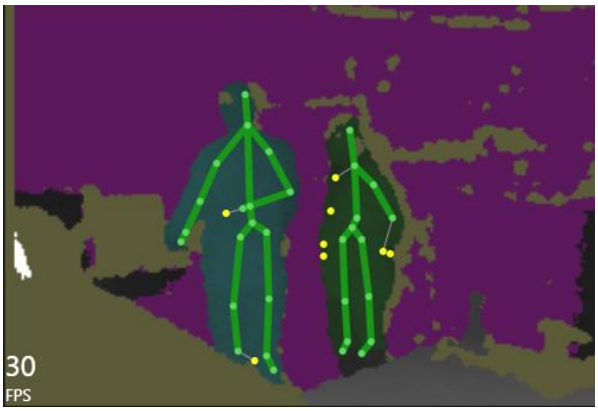


Figure 3. Skeletal tracking example.

B. Audio Localization Module

The microphone array of the Kinect is comprised of 4 supercardioid microphones that drive 24-bit ADC's. The frequency response of the microphones is tailored for human speech and their directivity is relatively stable for these frequencies (1-7 kHz). Sound source localization and beamforming are applied to the audio signal in order to determine the angle of the sound source in relation to the

device and acquire the audio signal from that particular direction (fig. 4). The returned values are the sound source angle in degrees in relation to the axis that is perpendicular to the device, and a confidence level of the reported angle.

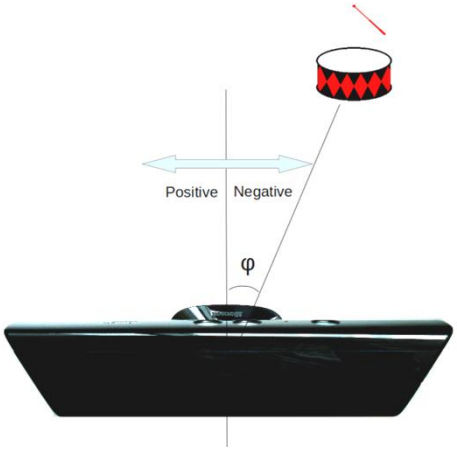


Figure 4. Kinect sound source localization.

Nevertheless, one Kinect is only capable of providing the angle of the sound source but not its distance, hampering localization accuracy. Therefore, we introduce a second Kinect to our system that is used solely for sound source localization (fig. 5). The second unit also provides an angle for the source of the sound, which can be used in combination with the previously obtained angle for accurate localization through triangulation. In order to do so, we need to obtain some data concerning the placement of the sensors. More specifically, let L be the distance between the two devices, A and B. Also, let θ_A, θ_B be the angle between the wall and the axis perpendicular to device A and B respectively. This angle should optimally be 45 degrees to maximize coverage assuming the devices are mounted at the corners of the same wall in a square room. Assuming there is a sound source S detected by the two devices, let the corresponding detected angles be $\phi_A, \phi_B \in (-50,50)$. These angles are positive when the sound source is estimated to be on the left side of the device and negative when the source is estimated to be on the right of the device (fig. 4). We will consider the triangle that is created, with A, S and B as its vertices. The altitude of the triangle that is passing from vertex S, divides L into a and b so that a+b=L. Let the length of the altitude (in our case the distance of the audio source/person from the wall) be X_s. Then, we can formulate the following equations:

$$\tan(\theta_A - \phi_A) = \frac{X_s}{a}$$

$$\tan(\theta_B + \phi_B) = \frac{X_s}{b}$$

Since L=a+b, the final solution to the system of equations is given by:

$$X_s = \frac{\tan(\theta_A - \phi_A) \cdot \tan(\theta_B + \phi_B) \cdot L}{\tan(\theta_A - \phi_A) + \tan(\theta_B + \phi_B)}$$

$$a = \frac{X_s}{\tan(\theta_A - \phi_A)}$$

$$b = \frac{X_s}{\tan(\theta_B + \phi_B)}$$

Thus, we can calculate the precise position of the audio source in the 2-D layout of the room.

Due to the nature of the sensor and propagation of sound waves some restrictions had to be imposed in order to ensure reliable location estimation. Therefore, the sound level is calculated for a window of 1 second and the sound source angles are taken into account only when the sound level exceeds 50dB, corresponding to a quiet conversation. This technique prevents inaccurate location estimation by ignoring low level background noise. Additionally, we only calculate the person's location when the confidence for both estimated sound source angles is more than 50%. A final and apparent restriction is that there must exist a solution for the equation system and this solution should fall within the monitored space. Thus, if a sound is coming from behind the sensors, or outside the limits of the monitored space, the location cannot be estimated or it is ignored respectively. This way, noises that are generated from external sources, e.g. a car passing-by, will not affect the location estimation.

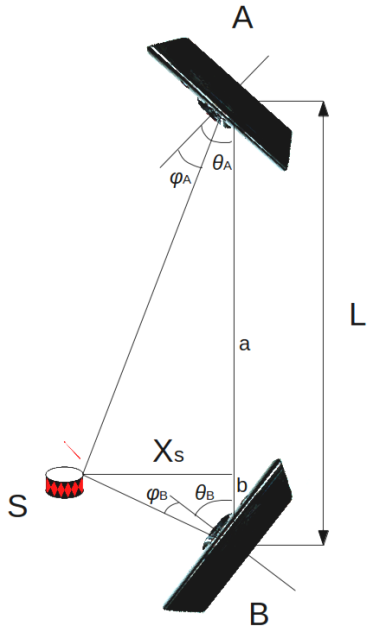


Figure 5. Configuration of the 2 Kinect devices for audio localization.

C. Automatic speech recognition module

In order to ensure natural interaction of the user with the system as well as effective emergency detection, we integrated an automatic speech recognition (ASR) module. This module is built using the hidden Markov model toolkit (HTK) [19]. The input to this module is the audio signal captured by the primary Kinect. The features extracted from the audio signal are 13 Mel frequency cepstral coefficients (MFCCs) and their first and second derivatives in order to account for speech

dynamics, creating a vector of 39 features. The models used are 3 state left-to-right HMMs modeling triphones. HTK is used for both training the models and recognizing speech. The module recognizes 11 words that comprise sentences commonly used in order to ask for help or assistance, e.g. "help me", "fire".

IV. System Operation

As mentioned earlier the two main functions of our system are person localization and emergency detection. These functions utilize information from both the skeletal tracking module and the audio localization module, but the ASR module is only used for emergency detection purposes. The following sections describe the two types of operation in more detail.

A. Person Localization

The main source of location information is the skeletal tracking module. More specifically, this module detects a person as soon as s/he enters the FOV of the sensor and tracks her/him while moving in the room. The accuracy and robustness of the tracker is exceptional due to the nature of the depth sensor, so the person is tracked while standing, walking or even sitting. We consider the location of the person as the average of the 3-D coordinates of all the tracked joints, expressed as $< \overline{X_d}, \overline{Y_d}, \overline{Z_d} >$, where:

$\overline{X_d} = \frac{1}{20}\sum_{i=1}^{20} X_{di}$ the mean distance from the sensor's plane.

$\overline{Y_d} = \frac{1}{20}\sum_{i=1}^{20} Y_{di}$ the mean deviation from the sensor's axis.

$\overline{Z_d} = \frac{1}{20}\sum_{i=1}^{20} Z_{di}$ the mean distance from the floor.

Another source of location information is the audio localization module. It should be noted that the audio localization module is capable of estimating the location of the person in 2 dimensions expressed by $< X_s$, a >, not accounting for height.

The final estimated location of the person is a result of combining the information from both modules. More specifically, when a location estimate is available from both modules, the average of each of the 2-D coordinates is calculated after proper transformation to match the 2 coordinate systems, while the third coordinate equals that of the skeletal tracking module. In the case where either of the modules does not return any coordinates, then the other module's coordinates are considered, e.g. if the person is outside the FOV of the depth sensor, only the audio localization coordinates are used. For our application, the detected activity is bound to the estimated location of the person. Therefore, if a person is standing by an appliance such as the oven or refrigerator we infer that s/he is using this particular appliance.

B. Emergency Detection

Our system is also capable of recognizing emergencies and due to the 3-D localization information it is very successful at detecting falls. In order to effectively carry out this operation, we utilize information from all three modules. In more detail, when coordinate $Z = \overline{Z_d}$ of the tracked skeleton falls below a predefined threshold (default is 2ft.), the system enters a stand-by mode. While in this mode, the system detects an emergency if any of the following conditions is met:

1. The ASR module recognizes that the person is asking for help.

2. A loud noise is recorded.

3. Z remains below the threshold and no sound is detected for a predefined period of time (default is 2 min.).

In addition, an emergency is detected even if the system is not in the stand-by mode when either of these conditions occurs:

1. The ASR module recognizes that the person is asking for help and skeletal tracking fails to locate the person.

2. The ASR module recognizes any of the predefined sentences for help repeated 3times in a 15 second window, independently of the status of the localization modules.

When any of the above 5 situations is identified, an emergency is detected and a distress signal is issued, including the person's last known location, in order to request for assistance.

V. Experimental Setup

An extensive set of evaluation experiments were conducted in order to fine-tune the parameters of the setup at our simulated assistive apartment (fig. 6). As mentioned earlier, two Kinect devices were used, mounted at the opposite sides of one of the walls, facing the entrance. The distance between the two devices was 175.5 inches. The axis perpendicular to the device points at 45 degrees towards the interior of the apartment, maximizing both the FOV and microphone coverage (fig. 7).

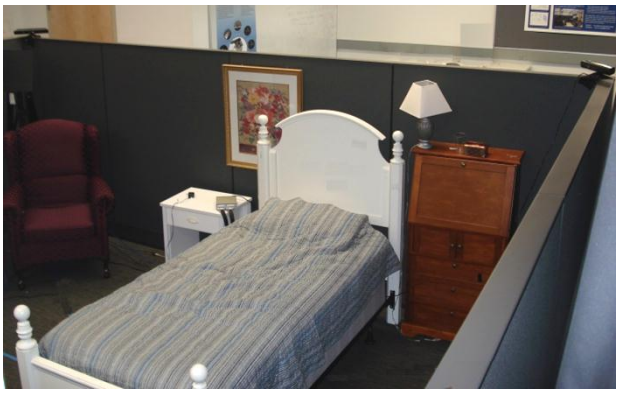


Figure 6. An aspect of our simulated assistive apartment.

All modules were installed on the same computer, although our system's implementation permits the use of separate computers for each one of the modules. For our experiments we partitioned the space in 8 different sectors, intersecting at the center of the room. The estimated location of the person was considered accurate when the coordinates fell within the boundaries of the corresponding sector. For our application, the detected activity is bound to the estimated sector.

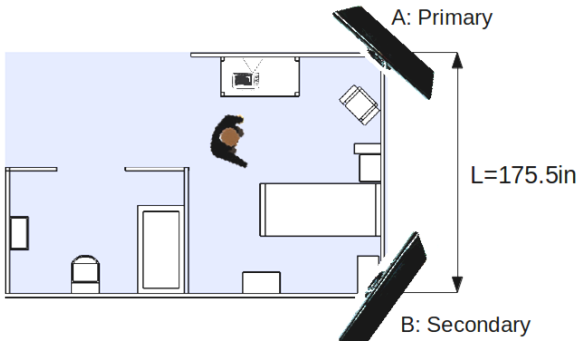


Figure 7. Simulated assistive apartment layout and placement of the Kinect devices.

Four individuals participated in our experiments, with either one or two occupying the apartment simultaneously. Subjects were asked to move in the apartment and perform every day activities. In the scenario where the space is inhabited by a single person, the location estimation of the person is always accurate as long as the person remains in the FOV of the sensor. Furthermore, in the event of 2 people being tracked, a realistic case where a caretaker or visitor is also present, the localization accuracy is 85% due to the people interacting and the resulting occlusions. In this case, person identification is not required, since we are primarily interested in the performed activities. Additional people may also be tracked but with reduced accuracy, however this is outside the scope of our system, since we aim at detecting activities and emergencies when the inhabitant is not supervised. Finally, the word accuracy of our ASR system was 94%, with low background noise levels.

VI. Conclusions

We presented a novel system capable of accurate and robust person localization and emergency detection. This system uses as input the depth sensor and microphone array of the Kinect device. Skeletal tracking and sound source localization are combined in order to estimate the position of the inhabitant. ASR is used as a natural means of interaction and in addition to the location information for emergency detection. The system was deployed in a simulated assistive environment and during the experiments conducted, it achieved both high localization and word recognition accuracy. After, confirming the effectiveness of our design we plan to extend it by utilizing depth information of additional Kinect devices for increased robustness and coverage.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grants No. NSF-CNS 1035913, NSF-CNS 0923494.

The authors would like to thank UTARI for its support.

References

[1] C. H. Kuo and R. Nevatia, "How does Person Identity Recognition Help Multi-Person Tracking?", *In Proc*. CVPR, pp. 1217-1224, 2011.

[2] D. Mitzel, E. Horbert, A. Ess and B. Leibe, "Multi-person tracking with sparse detection and continuous segmentation", *In Proc*. ECCV, pp. 397-410, 2010.

[3] M. Andriluka, S. Roth and B. Schiele, "People Tracking-by-Detection and People Detection by Tracking," In Proc. CVPR, pp. 1-8, 2008.

[4] B. Wu, R. Nevatia, "Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet Part Detectors", In Proc. IJCV, pp. 247–266, 2007.

[5] J. Krumm, S. Harris, B. Meyers, B. Brumitt, M. H. and S. Shafer, "Multi-camera multi-person tracking for easy living", In Proc. IEEE IWVS, pp. 3-10, 2000.

[6] D. Focken and R. Stiefelhagen, "Towards Vision-Based 3-D People Tracking in a Smart Room", In Proc. IEEE ICMI, pp. 400-405, 2002.

[7] A. Ess, B. Leibe, K. Schindler and L. V. Gool, "Robust Multi-Person Tracking from a Mobile Platform", In Proc. PAMI, pp. 1831–1846, 2009.

[8] V. M. Lee, T. W. Wong and C. C. Lau, "Home accidents in elderly patients presenting to an emergency department", Accident and Emergency Nursing, 7(2): 96–102, 1999.

[9] J. M. Hausdorff, D. A. Rios and H. K. Edelber, "Gait variability and fall risk in community–living older adults: a 1–year prospective study", Archives of Physical Medicine and Rehabilitation, 82(8): 1050–6, 2001.

[10] M. C. Hornbrook, V. J. Stevens, D. J. Wingfield, J. F. Hollis, M. R. Greenlick and M. G. Ory, "Preventing falls among community–dwelling older persons: results from a randomized trial", The Gerontologist, 34(1):16–23, 1994.

[11] J. A. Stevens, "Fatalities and injuries from falls among older adults", MMWR, 55(45), 2006.

[12] H. Nait-Charif and S. J. McKenna, "Activity summarisation and fall detection in a supportive home environment", In Proc. ICPR, vol.4, pp. 323- 326, 2004.

[13] C. Rougier, J. Meunier, A. St-Arnaud and J. Rousseau, "Fall Detection from Human Shape and Motion History Using Video Surveillance", In Proc. AINAW, vol.2, pp. 875-880, 2007.

[14] H. O. Alemdar, G. R. Yavuz, M. O. Ozen, Y. E. Kara, O. D. Incel, L. Akarum and C. Ersoy, "Multi-modal fall detection within the WeCare framework", In Proc. ICIPSN, pp. 436-437, 2010.

[15] "The PrimeSensor Reference Design", [online] available at: http://www.primesense.com/?p=514

[16] C. Liebe, C. Padgett, J. Chapsky, D. Wilson, K. Brown, S. Jerebets, H. Goldberg and J. Schroeder, "Spacecraft hazard avoidance utilizing structured light", In Proc. IEEE Aerospace Conference, pp.10, 2006.

[17] S. Rowe and C. Wagner, "An Introduction to the Joint Architecture for Unmanned Systems (JAUS)", Technical Report from Cybernet Systems Corporation, available at: http://www.cybernet.com

[18] "The Microsoft Kinect SDK", [online] available at: http://msdn.microsoft.com/en-us/library/hh855347.aspx

[19] S. Young, G. Evermann, D. Kershaw, G. Moore, J. Odell, D. Ollason, D. Povey, V. Valtchev and P. Woodland, "The HTK Book", Cambridge Univ. Eng. Dept., Tech Rep., 2002.

Authors Profile

Georgios Galatas is pursuing his PhD in Computer Engineering at the Heracleia Human-Centered Laboratory of the Computer Science and Engineering department of the University of Texas at Arlington. He is also a research fellow of the Institute of Informatics and Telecommunications of the National Center for Scientific Research "Demokritos". He received his combined bachelor's and master's degree from the Electrical and Computer Engineering department of the University of Patras in 2008. His research interests include Computer Vision, Digital Image Processing, Automatic Speech Recognition and Person Localization. He has co-authored several peer reviewed papers published in technical conferences and journals.

Shahina Ferdous received her PhD in Computer Science from the Department of Computer Science and Engineering of the University of Texas at Arlington in December 2012. She completed her B.Sc. from the Department of CSE of the Bangladesh University of Engineering and Technology (BUET) in 2006. She worked as a Software Engineer in Therap BD Ltd. during 2007-2008, until she joined the BS to

PhD program in 2008. She joined the Heracleia Human-Centered Computing Laboratory as a Research Assistant in 2009. Her research interests include Data Mining, Database, Sensor Network and Pervasive Assistive Applications. Dr. Ferdous has co-authored several peer reviewed papers published in technical conferences and journals. She has also served as a committee member and reviewer in many conferences.

Fillia Makedon is Distinguished Professor and Department Head of Computer Science and Engineering at the University of Texas at Arlington. She received her PhD in Computer Science from Northwestern University in 1982. Between 1991-2006, she was professor of computer science at Dartmouth College where she founded and directed the Dartmouth Experimental Visualization Laboratory (DEVLAB). Prof. Makedon has received many NSF research awards in the areas of trust management, data mining, parallel computing, visualization and knowledge management. She is author of over 300 peer-reviewed research publications. She directs the Heracleia Human-Centered Laboratory that develops assistive technologies for human monitoring and smart health. She is member of several journal editorial boards and chair of the annual PETRA conference (www.petrae.org).