Semantic Integration of Heterogeneous Recognition Systems

Paweł L. Kaczmarek and Piotr Raszkowski

Gdańsk University of Technology,
Faculty of Electronics, Telecommunications and Informatics,
Narutowicza 11/12 Str., 80-233 Gdańsk, Poland
{pawel.kaczmarek,piotr.raszkowski}@eti.pg.gda.pl
http://www.eti.pg.gda.pl

Abstract. Computer perception of real-life situations is performed using a variety of recognition techniques, including video-based computer vision, biometric systems, RFID devices and others. The proliferation of recognition modules enables development of complex systems by integration of existing components, analogously to the Service Oriented Architecture technology. In the paper, we propose a method that enables integration of information from existing modules to calculate results that are more accurate and complete. The method uses semantic description of concepts and reasoning to manage syntactic differences between information returned by modules. The semantic description is based on existing real-world concepts in video recognition and ubiquitous systems. We propose helper functionalities such as: module credibility rating, confidence level declaration and selection of communication protocol. Two integration modes are defined: voting of matching concepts and aggregation of complementing concepts.

Keywords: integration, ontology, computer perception, monitoring.

1 Introduction

Video-based computer vision is a typical method of computer perception of real-life situations. In a wider context, computer perception may be realized using biometric systems, RFID devices, environment condition sensors, and others. The techniques differ in the contents of input data and recognition algorithms, which determines their quality and the scope of application. Despite the differences, the systems have many similarities as all of them attempt to recognize concepts encountered in real-life.

The proliferation of computer perception systems [13] [2] [11] results in their overlapping functionality such as face recognition, people counting, car identification and others. Consequently, there exist many alternative components realizing similar functionality. It becomes possible to apply the Service Oriented Architecture (SOA) approach in this area and develop complex systems by integration of existing components. During the process, the developer selects from alternative components those that supply most desired attributes, for example low

C. San Martin and S.-W. Kim (Eds.): CIARP 2011, LNCS 7042, pp. 288–295, 2011.

price, high accuracy and high performance. The approach reduces development cost and time, but requires resolution of integration problems. The problems cover all layer of computer system, from the communication layer to semantic understanding of data [14] [6].

In the paper, we propose a method of integration of existing recognition modules in order to achieve results that are more accurate and more complete. We assume, that different modules supply information about the same situation, although the systems may recognize different elements of the situation and use incompatible descriptions. Our method integrates results from different sources and performs semantic reasoning to calculate a coherent description of a situation. We base our solution on ontological description of concepts that occur in the environment and reasoning rules that are applied to input data.

We propose two alternative integration modes in the method: voting and aggregation. The voting mode assumes that integrated modules recognize the same concept and the calculation aims at achieving a more accurate result. The aggregation mode, in turn, assumes that a complex information is composed from partial information returned by modules. Fig. 1 shows a concept diagram of integration by voting.

The rest of the paper is organized as follows. The next section describes related work and gives background about techniques used in our research. Sect. 3 describes main functionality of the solution. Sect. 4 describes system implementation. Finally, Sect. 5 concludes the paper.

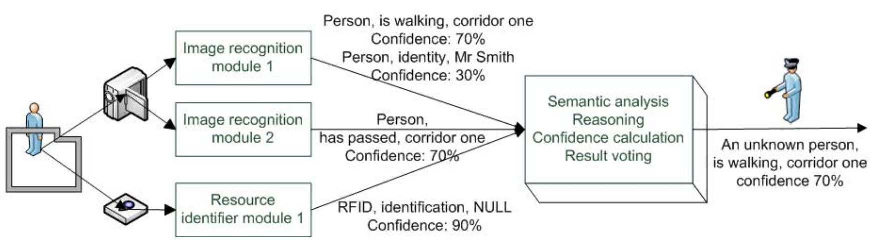

Fig. 1. General overview of module integration using semantic reasoning

2 Background and Related Work

Ontologies and semantic description of concepts are commonly used in computer vision systems in different aspects of the process. [8][5][10] propose ontology-based image retrieval, in which ontologies are used to describe both low-level visual properties, such as color, shape and texture, and high-level concepts regarding image contents, such as person, building and car. The SOUPA ontology [4] is a mature ontology designed to describe situations in ubiquitous and pervasive environments. The ontology contains concepts such as time, place, person, which largely overlaps with concepts encountered in image recognition systems.

[1] describes VERL (A Video Event Representation Language) - a formal language for describing an ontology of events that occur in real-life, and VEML (Video Event Markup Language) - used to annotate instances of the events described in VERL. The languages were designed as a part of the ARDA "Challenge Project on Video Event Taxonomy" project. The results supply formal mechanisms to describe events that can be identified during video analysis. [3] describes a similar approach, in which an ontology describes concepts related to video events. Two main types of concepts are distinguished: physical objects observed in a scene and video events occurring in a scene. In further work, we use selected concepts from existing ontologies to uniformly describe information retrieved by alternative modules.

Image retrieval is applied in a wide range of systems. For example, ubiquitous systems augment reality with computer-driven intelligence that automates every-day tasks and dynamically adapts to changing conditions [13]. The systems use devices and mechanisms of real-world perception such as move sensors, biometric systems, RFID devices. Monitoring systems use image retrieval to identify events that violate security restrictions and require appropriate actions. The systems use both video-based computer vision and non-video sensors for reality perception, analogously to ubiquitous environments [15] [9]. In our work, we propose a method of integration of existing real-world perception systems, rather than new methods of image and video processing.

Integration of components requires resolution of dependability and interoperability issues. The work [7] describes problems encountered during development of dependable applications out of undependable components. Authors propose a classification of component attributes and their rating. Typically, dependability is achieved using a variety of redundancy techniques [12]. The use of redundancy seems especially suited for SOA-based development, as there exist many modules supplying a similar functionality. Interoperability is another important issue in SOA-based applications. [6,14] overview existing definitions and metrics of interoperability. Metrics describe various levels of module integration, ranging from low-level communication protocol compatibility to high-level integration of information. Semantic understanding and ontological description of data concepts are important elements of high-level interoperability.

3 System Infrastructure and Operation

The proposed method assumes that there exist alternative computer perception modules supplying the same functionality. We integrate information from the alternative modules and calculate aggregated results to increase recognition accuracy and reliability. The method requires the following metadata information:

- A registry of integrated modules.
- An ontology of considered concepts.

The registry stores information about known recognition modules together with their credibility rating. Credibility ratings (denoted *cred*) are defined by administrator on module registration and adjusted during system operation. Initially, system operator defines relative rating of modules that are integrated with the system. During operation, the integration system monitors module results and calculates simple correlation between data received from modules. If the correlation of results for some modules fall below a specified threshold, the system informs administrator. The administrator is expected to adjust the ratings either manually or automatically by specifying feedback information about correctness of results returned from modules.

The ontology contains concepts encountered in real-life situations, analogously to ontologies like SOUPA [4], VERL [1] and WordNet. We anticipate confidence level (denoted $conf$) for input information and for output results, which allows application of fuzzy reasoning. Result confidence is calculated from confidence of input and credibility of the sending module as described in detail later.

3.1 Reasoning Rules

The system uses two alternative modes for integration of knowledge from recognition systems:

– voting mode,
– aggregation mode.

The voting mode attempts to detect a single concept and generate a result with higher accuracy. It is assumed that recognition systems supply alternative descriptions of the same situation, although they have limited functionality and may return imprecise results. For example, results from different face recognition systems are voted to determine the identity of a recognized person with higher accuracy.

The aggregation mode attempts to reason about a complex situation on the basis of detailed information. In this mode, we assume that integrated modules recognize the same scene, although they have complementing (rather than alternative) functionality. For example, alternative image recognition systems detect that a scene contains doors, windows and people, which enables us to reason that the scene contains a building in a public place. The concepts are aggregated and the system reasons about a possible complex situation.

The ontology is enriched with appropriate reasoning rules for both voting and aggregation modes. The voting mode requires mainly processing of the class structure to detect subclasses and superclasses of recognized concepts. The aggregation mode, in turn, requires rather analysis of concept relationships, such as "consists of", "contains", "stores".

Although modules should send information compatible with concepts defined in the common ontology, it may happen that unknown concepts are sent. In this case, the system applies a simple syntactic comparison of input data. In the approach, it is required that concepts supplied by different modules match exactly, that is if two modules recognize the same concept, they use the same word for description.

3.2 Information Processing Steps

Assuming that appropriate metadata has been configured and integration mode (voting, aggregation) has been selected, information processing in the system is done in the following main steps as shown in Algorithm 1.

Algorithm 1. Main steps of information processing in the integration system

input: data (d) sent from recognition modules ($R = (r_1, r_2, ..., r_N)$)
output: compound result

for all Received data **do**
  Create appropriate object structure for SPARQL processing.
  Search the ontology for received concepts.
  if Received concepts (in d) are found in the ontology **then**
    if There exist d_i, d_j that contain syntactically different data **then**
      Apply ontology-based reasoning
      Calculate sums of $cred * conf$ for inputs that are syntactically identical among R
      Select a reference input Ref as the input with highest value of $\sum cred*conf$
      Query ontology for semantic understanding of data (for example: Mr Smith, Mr John)
      Calculate a common base of concepts taking Ref as reference (for example: person)
    end if
    Apply voting weighted with relative conficence ($cred * conf$).
    Calculate confidence level of the result from module credibility and input data confidence.
    $Confidence = \frac{\sum_N cred*conf}{\sum_N cred}$
    Calculate correlation of results from modules to adjust credibility rating.
  else
    Received concepts are not matched
    Calculate relative confidence of each input $cred * conf$.
    Apply syntactic, majority voting weighted with relative confidence.
  end if
  Return compound result and confidence to the end user.
end for

4 System Implementation

As a part of the research, we develop a system that realizes the proposed method. The system supplies a web-based user interface that enables initial configuration and monitoring of operation. The current work covers registration of integrated recognition modules together with their credibility description as a major element of system configuration. Additionally, detailed configuration options are set, including, among others, specification of: integration modes (aggregation, voting), dictionary and ontology processing, time constraints for communication.

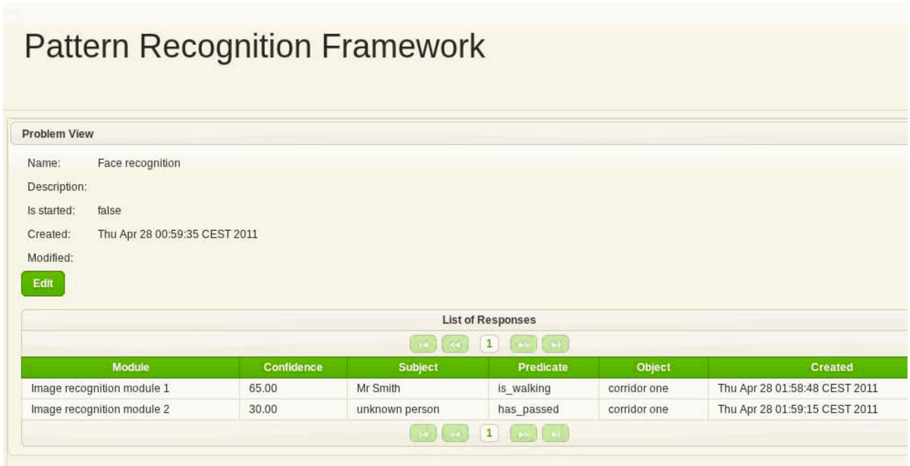

Fig. 2. Exemplary screenshot of the implemented system

The system is implemented in the Java 6 EE language using NetBeans IDE environment. We used the PrimeFaces Java framework to design a user-friendly web-based interface. Glassfish v3 Application Server is used as the deployment platform for the system MySQL 5 database is used as persistent storage. Fig. 2 shows an exemplary screenshot of the implemented system.

4.1 Ontology Processing in the System

We use Protégé as the editor for the ontology stored in the system. Our work focuses on defining concepts related to office area. This includes classes such as person, room, action, device, and appropriate individuals, for example Mr Smith, Mr Jones, Room100, Room200 etc.

Runtime ontology processing is performed using Jena OWL API and the Sparql processing engine. Sparql queries aim at retrieving information from the ontology that will be useful for common understanding of received data. As as example, consider the figure from the Introduction section (Fig. 1). The first module detected that a person is walking (confidence 70%) and the person is Mr Smith (confidence 30%). The second module detected that a person has passed (confidence 70%). The RFID system informs that is has not received any identification (confidence 90%). Therefore, we conclude that someone is walking the area, but it is not necessarily Mr Smith. Therefore, the reasoning should return information that an unknown person is walking the area, while the confidence of the information is 70%, assuming that modules have equal credibility.

4.2 Communication Protocol

The system requires that recognition modules send knowledge organized analogously to N-triples, containing the subject, the predicate and the object.

The communication format enables relatively simple processing on the recognition module side. The communication protocol anticipates grouping of triples into one logical set. In this case, a module needs to send a control triple informing that following communication should be merged into one information. Additionally, confidence level may be assigned to transmitted information.

Two concrete communication interfaces are supplied: the Web services interface and the socket interface. The Web services interface defines the sendTriple method for single communication and the registerResponse method for grouping following communication into one information. sendExtendedTriple enriches data with the assigned confidence level. The socket interface supplies analogous functionality using a lower level communication mechanism. The interface enables transmission of integer operation codes and character arrays of information, for example: 31 - start of triple, 32 - end of triple, 41 - start of extended triple. The interface is anticipated for systems that focus on low-level solutions and are difficult to integrate with Web services communication libraries.

5 Conclusions and Future Work

The proposed method intends to integrate information from independent recognition modules in order to achieve more accurate and complete results. The approach is driven by the SOA technology, in which applications are developed from existing, alternative modules. The use of open communication standards will promote interoperability and easy integration of modules. The current implementation work enables us to refine and adjust the method to concrete cases.

Integration of concrete recognition modules will be the main scope of future work. We plan to integrate both our proprietary implementations and existing modules, which requires minor adjustment of modules to integration system requirements. The adjustment covers two main areas: (i) implementation of appropriate network interface for data transmission and (ii) use of common concepts for description of recognized elements. During method design, we intended to minimize the work that is required to integrate recognition modules.

Adjustment of reasoning rules and ontological description of concepts is another interesting area of future work, as currently we analyze a limited number of rules and concepts. Extension of the knowledge base enables application of the system in a wide range of areas. Existing knowledge bases, such as WordNet or SOUPA ontology, supply virtually unlimited possibilities of concept definition and processing. We hope that the method will increase accuracy of recognition systems in the future and will promote application of existing systems in new areas.

Acknowledgments. The work was supported in part by the Polish Ministry of Science and Higher Education under research projects N N519 172337 and N N519 432338.

References

1. Bolles, B., Nevatia, R.: A hierarchical video event ontology in owl. Tech. rep., 2004 ARDA Challenge Project (2004)
2. Bradski, G., Kaehler, A.: Learning OpenCV: Computer Vision with the OpenCV Library. O'Reilly Media (2008)
3. Bremond, F., Maillot, N., Thonnat, M., Vu, V.T.: Ontologies for video events. Tech. rep., Institut National de Recherche en Informatique et en Automatique (2004)
4. Chen, H., Perich, F., Finin, T., Joshi, A.: Soupa: Standard ontology for ubiquitous and pervasive applications. In: International Conference on Mobile and Ubiquitous Systems: Networking and Services (2004)
5. Dasiopoulou, S., Papastathis, V.K., Mezaris, V., Kompatsiaris, I., Strintzis, M.G.: An ontology framework for knowledge-assisted semantic video analysis and annotation. In: 3rd International Semantic Web Conference (2004)
6. Ford, T., Colombi, J., Graham, S., Jacques, D.: A survey on interoperability measurement. In: 12th ICCRTS Adapting C2 to the 21st Century
7. Gorbenko, A., Kharchenko, V., Popov, P., Romanovsky, A., Boyarchuk, A.: Development of dependable web services out of undependable web components. Tech. Rep. CS-TR-863, University of Newcastle upon Tyne (2004)
8. Hois, J., Wünstel, M., Bateman, J., Röfer, T.: Dialog-based 3d-image recognition using a domain ontology. LNCS (2007)
9. Kaczmarek, P.L., Zielonka, P.: A video monitoring system using ontology-driven identification of threats. In: International Conference on Human System Interaction (2009)
10. Maillot, N., Thonnat, M., Boucher, A.: Towards ontology-based cognitive vision. Machine Vision and Applications, Special issue on ICVS (2003)
11. Szeliski, R.: Computer Vision: Algorithms and Applications. Springer, Heidelberg (2010), `http://szeliski.org/Book/`
12. ReSIST: Resilience for Survivability in IST, A European Network of Excellence: Resilience-Building Technologies: State of Knowledge (2006)
13. Poslad, S.: Ubiquitous Computing: Smart Devices, Environments and Interactions. Wiley Publishing (2009)
14. Tolk, A.: Beyond technical interoperability - introducing a reference model for measures of merit for coalition interoperability. In: 8th International Command and Control Research and Technology Symposium, Washington, D.C. (2003)
15. Tyrer, J.: Dtect - the vehicle monitoring system. Loughborough University, Vehicle Occupancy Ltd (2007)