Multimodal Integration (Image and Text) Using Ontology Alignment

Ahmad Adel Abu Shareha, Mandava Rajeswari and Dhanesh Ramachandram
School of Computer Sciences, University Sains Malaysia, 11800 Penang, Malaysia

Abstract: Problem statement: This study proposed multimodal integration method at the concept level to investigate information from multimodalities. The multimodal data was represented as two separate lists of concepts which were extracted from images and its related text. The concepts extracted from image analysis are often ambiguous, while the concepts extracted from text processing could be sense-ambiguous. The major problems that face the integration of the underlying modalities (image and text) were: The difference in the coverage and the difference in the granularity level. **Approach:** This study proposed a novel application using ontology alignment to unify the underlying ontologies. The said lists of concepts were represented in a structured form within the corresponding ontologies then the two structural lists are enriched and matched based on the alignment, this matching represent the final knowledge. **Results:** The difference in the coverage was solved in this study using the alignment process and the difference in the granularity level was solved using the enrichment process. Thus, the proposed integration produced accurate integrated results. **Conclusion:** Thus, integration of these concepts allows the totality of the knowledge be expressed more precisely.

Key words: Concept-level multimodal integration, ontology alignment and semantic knowledge

INTRODUCTION

Multimodal fusion and multimodal integration refers to the merging of different sources of information as a means of enhancing the outcome some specific task. The richness of the information provided by multimodal data could potentially lead to better performance than those tasks that rely on unimodal data. In the natural world, human and animals in the higher rungs of evolution perceive the world using multiple senses concurrently and use their acquired knowledge to analyze and to understand events. Here, multimodal information integration takes place at a high level using predetermined knowledge[1,2]. Data used in multimodal integration can be present in different levels of abstraction.

There have been several of multimodal integration approaches reported in the literature that vary in their context and application. Generally, we may categorize these approaches into (a): Multimodal Fusion Approaches, (b): Tightly-coupled Multimodal Integration and (c): Augmented Unimodal Analysis. In multimodal fusion, low-level integration of multimodal data is the main characteristic of the approach, rich data from a single source is divided into multiple modals for efficient processing and finally combined for better interpretation[3,4]. Tightly-coupled multimodal integration involves data from multiple sources that are tightly coupled (e.g., movements of the lips to words that are being read in the speech) which is processed independently and integrated at a high level of abstraction. If we consider this example, both the image and text express the same information at any given time[5,6]. The multimodal integration is then performed at a higher level of data abstraction. In, what we classify as augmented unimodal analysis, the extraction of knowledge is primarily based on a dominant modality. However, to aid the analysis and the interpretation of the subject matter of interest, associated data from a different modality may be used. Here, the assisted knowledge is used without any pre-processing, hence the data from dominant modality has to be processed and transformed to a form suitable to be used with the assisting modality. For example, in the research of Benitez and Chang[7], perceptual knowledge is extracted from an image and then disambiguated with the assistance of the associated keywords. Here, the main focus is to disambiguate image content using textual data (keywords).

In the context of our research, the multimodal information consists of images and the accompanying textual descriptions of the images or any free text

Corresponding Author: Ahmad Adel Abu Shareha, School of Computer Science, University Sains Malaysia, 11800, Penang, Malaysia Tel: +604-6533888 Fax: +604-6573335

related with the image. We are proposing a methodology to integrate this multimodal information at the semantic concept level. Hence, the term concept-level multimodal integration.

It is highly possible that the concepts extracted from both image as well as text may be ambiguous. However, we assume that the concepts extracted from the associated text data are less ambiguous than those extracted from the image. This assumption is justified from the fact that presently, research into concept extraction from textual data is at a more advanced stage than it is for image data. Thus, our study leans towards disambiguation of image concepts using associated text concepts.

The semantic concepts that we use here are assumed to be readily available or whose prior extraction is performed using some ontology-based methods. In many cases, the concepts extracted from images are often ambiguous, while the concepts extracted from text analysis are often sense-ambiguous. Singly, each of these concepts cannot express in totality the sum of knowledge contained in the data source as opposed to the richness of knowledge that can be expressed as a result of their integration. The major problems that face the integration of the underlying modalities (image and text) are: The difference in the granularity level which is solved in this study using the alignment process and the difference in the coverage which is solved using the enrichment process. As ontology-based image and text processing are prerequisites for the framework proposed in this study, the reader is referred to[3,8,9] for some background on ontology-based text processing. We focus on the technique of integrating the ontologies, or in a simpler data structure, lists of concepts.

Ontologies are fundamental to our research; therefore we briefly introduce them here. In both computer science and information science, an ontology is a data model that represents a set of concepts within a domain and the relationships between those concepts. It is used to reason about the objects within that domain. Ontologies are used in artificial intelligence, the Semantic Web, software engineering, biomedical informatics and information architecture as a form of knowledge representation about the world or some parts of it[10,11]. Ontology alignment is the problem of finding similarities between two or more ontologies. Historically, ontology alignment was sought to integrate heterogeneous databases which were independently developed and contained differing vocabularies. The research focus was to find "semantically equivalent" classes of data. There are several means to achieving ontology alignment-one approach is known as lexical

methods[12] in which the abstract data (nodes) represented by its names are used for alignment, while the second manner in which ontology alignment can be performed is through structural methods[13-18]. This approach exploits the relationships (edges) between the different classes which form the structure of the ontology to facilitate the alignment.

Ontology alignment has been an important research topic in recent years. Since there are various groups of researchers developing ontologies for the same domain of knowledge, these ontologies often heterogeneous. Hence, there is a need for ontology alignment to unify or enrich the body of knowledge. In computer science, ontologies have been extensively used by computational linguists. A classic example where ontology alignment has been used is in computer-aided translation. Language databases are in fact huge ontologies that represent the lexical and the structural characteristic of the languages (e.g., WordNet[19] for English language). Computer-aided translation of different languages has been made possible by the alignment of their respective ontologies[20]. Word sense disambiguation[21,22], which is the problem of finding the true sense of the word within the context in either monolingual or multilingual text has been solved using the ontology alignment techniques.

Another emerging area that requires ontology alignment is the Semantic Web. The Knowledge Web Consortium has published a technical report describing the state of art in ontology alignment for the Semantic Web[23]. While a number of techniques for ontology alignment have been presented, the authors conclude that very few comparisons or integrations have been implemented thus far. Towards this end, another research group called the Ontology Alignment Evaluation Initiative[24] has worked on evaluation strategies for ontology alignment techniques that could be employed to improve on current techniques.

In the traditional web application context, Fossati *et al.*[18] investigated the similarities between these various search engine directories. Web agents that used these different search engines could understand each other through the result of the alignment of their respective ontologies[25]. Ontology alignment has also been used to extend the capabilities of mobile device by enabling interoperability between different makes[26].

Ontology alignment has also been used for bioinformatic ontologies. Bio-ontologies and databases that contain information about genes, gene sequence information, proteins, gene functions, pathway information, genetic diseases, phenotypes and others contain knowledge that is inter-related and researchers would benefit if it were better connected.

This study proposes a novel application scenario of ontology alignment in the case of multimodal data integration involving images and related text. We discuss the approach in the following.

Concept extraction from image has been proven to be extremely challenging. The most straightforward manner in which "concepts" can be defined in images are through low-level image features. An aggregate of low-level image features may be associated to a particular high-level semantic concept. In reality, many high-level semantic concepts may share very similar low-level image features. As a means to resolve this ambiguity, this study proposes to exploit concepts extracted from related textual description which may be present for a given image.

The concepts that are identified from image and text processing methods may be incorrect or incomplete. As the text analysis and image analysis are performed independently from each other, it is possible that points of inconsistencies in the concept identification could be different. In event that the concept identification of, say, the textual domain is very accurate and that of the image domain is rather poor, then the concrete knowledge gained from the textual analysis may be used to improve the deficiencies of the image analysis. The integration of those two modals will produce a more accurate and better-structured knowledge.

The framework adapted in this research to integrate the knowledge generated from the image and its associated text consists of four major processes: (i) Ontology alignment (ii) Concept mapping and enrichment (iii) Graph matching and (iv) Building the final knowledge. Each of these processes is described in detail in the following subsections.

Ontology alignment: Ontology alignment is the problem of finding similarities between two or more input ontologies, by identifying the identical elements over the input ontologies. Two main approaches have been discussed in the literature: one which exploits the abstract data (nodes) represented by its names (lexical methods)[12] and another which exploits the relationships (edges) between the various classes that form the structure of the ontology, (structural methods)[13-18].

In present study, as the structure of the ontologies are expected to be slightly different, ontology alignment performed using a two-stage approach, initially using a lexical method first and followed by a structural approach. In the first stage, a simple string comparison method (edit distance) is used to initialize a number of similarity points between the underlying ontologies.

Then, a structural alignment method is performed based on the initial matches found prior. The structural method considers the following criteria to set up similarity between a pair of nodes:

- All or some of their direct super-entities are similar
- All or some of their sibling entities are similar
- All or some of their sub-entities are similar
- All or some of their descendant entities (entities in the sub tree rooted at that entity) are similar
- All or some of their leaf entities are similar
- All the entities in the path from the root to the correspondence entities are similar
- All of their relative entities are similar[27]

The ontology alignment results enable us to perform matching between the knowledge extracted from image and that extracted from text. In comparison, previous approaches[7] perform matching between the both forms directly. These approaches are limited by ability of the pre-processing stage in identifying the same level of data abstraction. Thus, such approach is not able to accurately process some input modalities (image and text) which have different granularity and coverage levels. In addition, previous research were dealt with unimodal information as well as similarity in amount and details of information. Figure 1 shows the ontology matching result of two input ontologies.

Mapping concepts to ontology and concepts enrichment: As mentioned earlier, image and text are processed independently to extract pertinent concepts in both modalities. In this process, the extracted concepts are independently mapped to the domain ontology, thus producing two graphs; one for each modality.
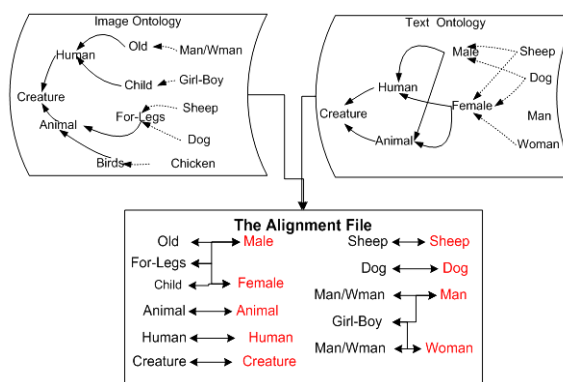

Fig. 1: Ontology alignment for image ontology and text ontology

The components in each of these graphs are the extracted concepts mapped into the ontology and connected to each other via the parent nodes (only). The connected concepts of a given document or image is referred to as "graph". This graph represents the complete knowledge of the concepts in the corresponding modality. The formation of the graph may be seen as an enrichment process due to its ability to add additional knowledge about the identified concepts. Limiting the connection via the parents ensures that concepts irrelevant to the image or the text document are not added on into the graph.

Concept enrichment is carried out on all concepts starting from those at lower level by adding parents' concepts which are at higher level. The enrichment continues, adding concepts at each subsequent levels sequentially, on until reaching a point where all the concepts have one sharing root. As an example, if the concept identification process for the image or the text process resulted in two sibling concepts, then the enrichment process combines those two concepts by adding the corresponding parent concept which leads to formation of a single graph of the two concepts and their parent. The enrichment process continues by adding the super classes sequentially which will lead finally to the combination of all the concepts for a given modality in a single graph.

Connecting the concepts to the parents (enrichment) helps to clarify concepts further through the integration of multiple modalities. In the case that the text is general and gives a high description of the image, the concepts from text cannot be matched with the concepts from the image. This is because there exists no node-to-node correspondence through the ontology alignment. The same case will occur when the text provides information at finer detail while image processing can only identify the general category of some objects. The enrichment process solves the problem of granularity in the modalities by linking the concepts with its higher level concept sequentially. Figure 2 shows an example of concepts mapped to ontology and enriched with parent nodes.
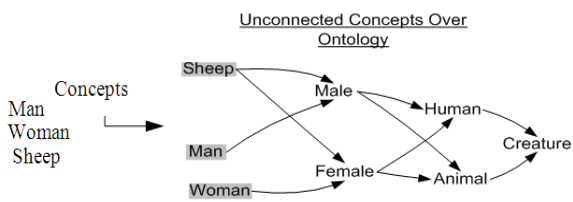

Fig. 2: Mapping and enrichment

Graph matching through ontology alignment: This process consists of identifying identical concepts in both modalities. Modality matching is achieved using the alignment result; if the pre-defined alignment process has identified the identical concepts in both ontologies, the identical concepts in the ontologies are identical in the graphs as well, since the constructed graphs are the sub-graph of the ontologies themselves.

Since the alignment is not visible to all the concepts in the ontologies or it might not be correct for all them, thus for small number of concepts from both ontologies it might be no alignment at all, the added parents helped in offering more concepts to the matching process and more over the concert knowledge represent in the high level parents is more properly to be matched than single objects (leaves).

The matching process identifies identical knowledge (trees, sub-trees, or branches) in both modalities, thus the ambiguous objects will be disambiguated by presenting the relevant knowledge and discarding the irrelevant knowledge in one modality which has no correspondence in the other modality.

Normally, in the case of multi-object image whose identified concepts are ambiguous and general in nature, the image will be represent in a single strongly connected graph which might include relevant and irrelevant knowledge (the image analysis is able to identify all the possible concepts for an ambiguous object), each possible concept will form a branch or sub-graph in the enrichment process. The text might form a larger graph because it includes more information than the text. In such case, the extracted concepts from the text will exceed in number the extracted concepts from the image which later form a large graph from the text side.

As in the case of image concepts, the sense-ambiguous words extracted from the text will form different branches. The matching process would eventually connect identical sub-graphs from each modality and discard the irrelevant nodes which represent the wrongly identified concepts and sense from image and text.

After enrichment and disambiguation is complete, the final knowledge is structured and connects concepts that represent the knowledge integration in both image and its corresponding text.

Output construction: The pairs of identical output nodes which are extracted from image and text and integrated through ontology alignment, are rearranged in a graph based on the relationships in the original ontologies. The sub-class and super-class relationships are reconstructed again based on the text ontology.
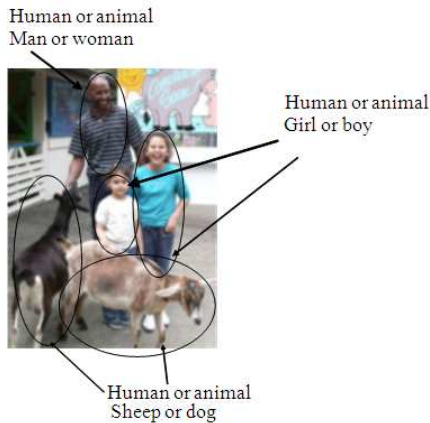
Fig. 3: Examples of image and related text (the picture shows the kindness relationship between the human and the animal. Girls with their parent feed sheep)

The result is a graph of knowledge that represents the final text is general. Thus, the text graphs include unrelated information which has to be eliminated. For the image, since few objects from different sub-domain might share the same low-level features, this similarity might lead to the construction of sub-graphs in the image knowledge. The benefit of the enrichment is evident in this final knowledge as it allows for further disambiguation of the original ambiguous objects, if needed.

Example: The inputs to the multimodal integration are two sets of concepts. The two set of concepts are mapped into graphs. For the text, the graph is big if the graph that corresponds to different sub-domains and different interpretation of the image.

Accuracy of the results depends heavily on the following factors: The relevancy of the associated text, the ambiguity of the objects, the lesser the ambiguity of an object, the easier it is for the disambiguation process and the domain of interest which has to be well-defined and the correspondence ontology has to be well-structured in order to get an accurate results.

Let us consider the example in Fig. 3, based on our experience in image analysis, we may, to some degree of confidence, assume that the concepts extracted from the image are: person and animal. Any further details about these concepts may either require extensive processing or simply may not be possible. On the other hand an examination of the associated text reveals that we may be able to retrieve concepts such as parent, girls and sheep that are relevant to the concepts from the image. Thus the knowledge from the text may be utilized to gain additional knowledge from the image.
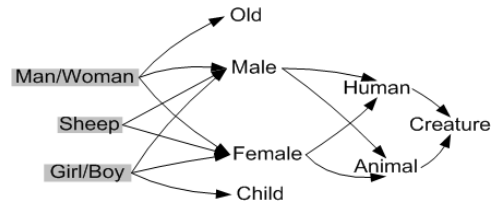

Fig. 4: The output for the example in Fig. 3

In addition, there may also be instances where image analysis may misidentify concepts. In the sample image, based on the shape feature, an object may be "animal" or a car. This ambiguity may be resolved using the text. Figure 5 shows the steps followed to extract knowledge from the image and text in Fig. 3.

The output knowledge from this example is getting as shown in Fig. 4.

MATERIALS AND METHODS

The proposed framework was implemented in Java using Alignment API, Protégé API and Jena API.

We have implemented and tested the proposed framework using simulated data. In the image part, 50 images first were collected from PASCAL dataset[28]. The collected images belong to 'Farm', 'Animal' and 'Human' domains. Then, the objects in these images were labeled manually with multi-labels include the real object label and similar objects with the same low level features. The image objects were totally ambiguous with 1, 2 and 4 concepts to test the ability of the proposed system with different rates of ambiguity in the image. The text was collected from the internet using keywords that selected from the image concepts. The long text was cut to 100 length words. Each image was associated with 3 different texts which classified into: irrelevant text, relevant and strongly relevant. The text concepts were extracted manually. Only nouns and noun phrase were selected to represent the text concepts.

Ontologies from FOAM[29] alignment tools were selected and modified. The original ontologies by FOAM are provided as dataset for alignment tests.

The true integrated outputs were constructed manually to be the ground truth for the output results. The ground truths for the presented cases with the irrelevant text are empty sets. The ground truths for the cases with strongly relevant text are richer or similar to the ground truth for the corresponding cases with relevant text. The corresponding cases with different degree of image ambiguity (1, 2 and 4) have similar ground truth.
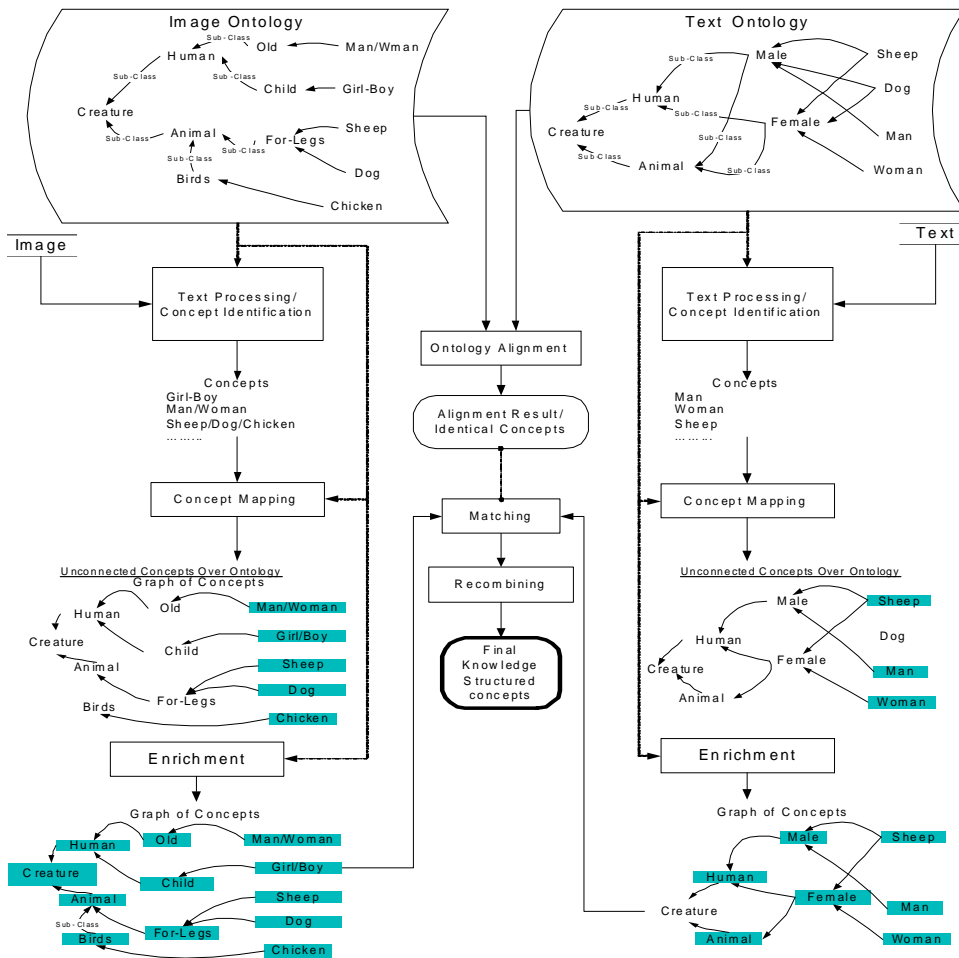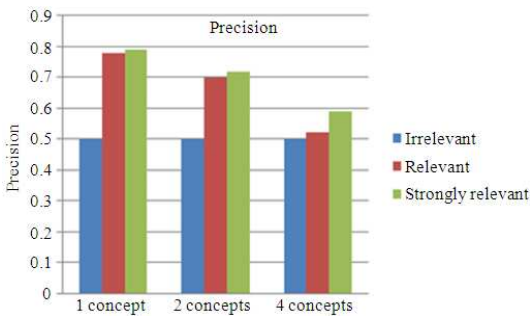
Fig. 5: Multimodal integration (image and text)

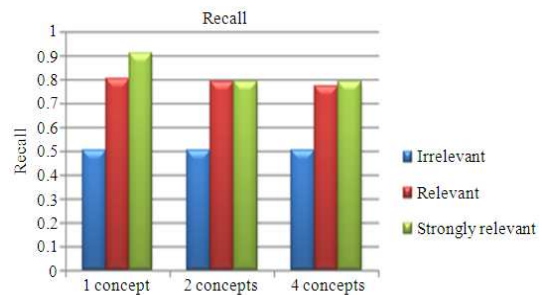Fig. 6: The precision of the integrated results

Fig. 7: The recall of the integrated results

RESULTS

The output results have been evaluated using the well-know measurements precision and recall and by referring to the ground truths provided for each case. Precision is the ratio between the true positive concepts retrieved by the method to the total number of the retrieved concepts. Recall is the ratio between the true positive concepts retrieved by the method to the total number of the true concepts in the corresponding ground truth.

The average precision and recall for the nine different scenarios (three degrees of images' concepts ambiguity and three text types) are shown in Fig. 6 and 7.

DISCUSION

As stated early the output results are strongly related to the relevancy of text. The recall value for the cases with strongly relvant text are better than those with relavant text. However, with mudium degree of relvency text, the integration has resulted in a good output. Overall, the results reveal a high precision because the implmented approach filleterd out the concepts that stated (with different level of abstraction) in both modalities. Yet, in some cases with highly ambiguated image concepts, the text might be integrateed wrongly with unreal image concepts (if the real and unreal image concepts fall under some tiny branch in the underlaying ontology) in a higher level of abstraction which reduce the precision of the integration output. The recall and the precision for the cases of the irrelevant text are fair since the integrated framework try to maximize the integration of the underlying modalities which is such cases caused false positive.

As a conclusion, the proposed approach has achieved a good precision in integrated image and text with different level of abstraction.

CONCLUSION

In this study, we have shown how the lists of concepts from the image and the text (multimodal data) can be integrated though the ontology alignment. The goal of this integration is to produce a semantic knowledge that is more rich and accurate than the knowledge that can be produce from any of them individually. The proposed study is novel method of multimodal integration that implements the integration at semantic concept level. The two concepts' lists can be obtained from any ontology-based image and text processing approaches. Graph is build from each data type (image and text) by using mapping and enrichment method; those clusters are then matched by referring to the ontology alignment.

The proposed approach treats the different modalities independently with the same priority, in which case the points of inconsistencies might appear in one modality and over come by the other. The integration is knowledge based, which is based on the agreement of the knowledge sources in the so called ontology alignment. This can be called knowledge-based integration, which set the matching between the modalities more precisely.

ACKNOWLEDGMENT

REFERENCES

1. Wu, Y., E.Y. Chang, K.C.C. Chang and J.R. Smith, 2004. Optimal multimodal fusion for multimedia data analysis. Proceedings of the 12th Annual ACM International Conference on Multimedia, Oct. 10-16, New York, USA., pp: 572-579. DOI: http://doi.acm.org/10.1145/1027527.1027665

2. Papageorgiou, H. and A. Protopapas, 2003. CIMWOS: A multimedia, multimodal and multilingual indexing and retrieval system. Proceedings of the 4th European Workshop on Image Analysis for Multimedia Interactive Services, (EWIAMIS'03), Queen Mary, University of London, World Scientific, Singapore, pp: 563-568. http://www.xanthi.ilsp.gr/cimwos/pdf/CIMWOS-WIAMIS.pdf

3. Tiun, S., R. Abdullah and T.E. Kong, 2001. Automatic topic identification using ontology hierarchy. Proceedings of the 2nd International Conference on Computational Linguistics and Intelligent Text Processing, (ICCLITP'01), Springer-Verlag, Malaysia, pp: 444-453. http://utmk.cs.usm.my/pdf/CICLing01.pdf

4. Cees, G.M. *et al.*, 2005. Early versus late fusion in semantic video analysis. Proceedings of the 13th Annual ACM International Conference on Multimedia, Nov. 6-10, Hilton, Singapore, pp: 399-402. DOI: http://doi.acm.org/10.1145/1101149.1101236

5. Lizhong, W., S. Oviatt and P. Cohen, 1999. Multimodal integration: A statistical view. IEEE Trans. Multimedia, 1: 334-342. DOI: http://doi/10.1109/6046.807953

6. Adler, A. and R. Davis, 2006. Speech and sketching for multimodal design. Proceeding of the International Conference on Computer Graphics and Interactive Techniques, July 30-Aug. 3, Boston, Massachusetts, pp: 214-216. DOI: http://doi.acm.org/10.1145/1185657.1185792

7. Benitez, A.B. and C. Shih-Fu, 2002. Perceptual knowledge construction from annotated image collections. Proceedings of IEEE International Conference on Multimedia and Expo, (ICME'02), Lausanne, Switzerland, pp: 189-192. DOI: 10.1109/ICME.2002.1035750

8. Berka, P., T. Athanasiadis and Y.S. Avrithis, 2006. Rule-based reasoning for semantic image segmentation and interpretation. Proceedings of the 1st International Conference on Semantics and Digital Media Technology, (ICSDMT'06), Athens, Greece, pp: 39-39. http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-233/p39.pdf

9. Grover, C., H. Halpin, E. Klein, J.L. Leidner, S. Potter, S. Riedel, S. Scrutchin and R. Tobin, 2004. A framework for text mining services. Proceedings of the 3rd UK e-Science Program All Hands Meeting, (SPAHM'04), UK., pp: 67. http://citeseerx.ist.psu.edu/viewdoc/summary?doi= 10.1.1.60.5806

10. Ontology Works, Inc., 2007. Ontology Development and Consulting. http://www.ontologyworks.com/

11. W3C World Wide Web, 2004. OWL Web Ontology Language. http://www.w3.org/TR/owl-features/

12. Li, J., 2004. LOM: A lexicon-based ontology mapping tool. Proceedings of the Performance Metrics for Intelligent Systems, (PMIS'04), Palo, Alto, pp: 1-5. http://www.isd.mel.nist.gov/research_areas/researc h_engineering/Performance_Metrics/PerMIS_2004 /Proceedings/Li%20final.pdf

13. Euzenat, J. and P. Valtchev, 2003. An integrative proximity measure for ontology alignment. Proceedings of the 1st International Workshop on Semantic Integration, (IWSI'03), pp: 1-6. http://www.dit.unitn.it/~accord/RelatedWork/Matc hing/euzenat2003h.pdf

14. Euzenat, J. and P. Valtchev, 2004. Similarity-based ontology alignment in OWL-Lite. Valencia, España. http://www.iro.umontreal.ca/~owlola/pdf/align-ECAI04-FSub.pdf

15. Kotis, K., A. Valarakos and G. Vouros, 2006. AUTOMS: Automated ontology mapping through synthesis of methods. Proceeding of the International Workshop on Ontology Matching Collocated with the 5th International Semantic Web Conference, (WOMCISWC'06), Samos, Greece, pp: 73-95. www.dit.unitn.it/~p2p/OM-2006/8-automs-OAEI'06.pdf

16. Fridman, N.N. and M.A. Musen, 2000. PROMPT: Algorithm and tool for automated ontology merging and alignment. Proceedings of the 17th National Conference on Artificial Intelligence and 12th Conference on Innovative Applications of Artificial Intelligence, (NCAICIAAI'00), AAAI Press/The MIT Press, USA., pp: 450-455. http://dit.unitn.it/~p2p/RelatedWork/Matching/SMI -2000-0831.pdf

17. Pan, R., Z.L. Ding, Y. Yu and Y. Peng, 2005. A Bayesian network approach to ontology mapping. Semantic Web-Iswc 2005. Lecture Notes Comput. Sci. 3729: 563-577. DOI: http://dx.doi.org/10.1007/11574620_41

18. Fossati, D., G. Ghidoni, B.D. Eugenio, I. Cruz, H. Xiao and R. Subba, 2006. The problem of ontology alignment on the web: A first report. Proceedings of the 11th Conference of the European Association of Computational Linguistics on Web as Corpus, EACL-Italy, pp: 51-59.

19. Fellbaum, C., 1998. WordNet: An Electronic Lexical Database. 2nd Edn., MIT Press, ISBN: 026206197X, pp: 423.

20. Ressler, J., M. Dean, E. Benson, E. Dorner and C. Morris, 2007. Application of ontology translation. Lecture Notes Comput. Sci., 4825: 830-842. DOI: 10.1007/978-3-540-76298-0_60

21. Ide, N. and J. Veronis, 1998. Word sense disambiguation: The state of the art. Comput. Linguist., 24: 1. http://www.up.univ-mrs.fr/~veronis/pdf/1998wsd.pdf

22. Ion, R. and D. Tufis, 2004. Multilingual word sense disambiguation using aligned WordNets. Romanian J. Infom. Sci. Tech., 7: 183-200. http://www.ceid.upatras.gr/Balkanet/journal/17_Se manticValidation.pdf

23. Euzenat, J., 2004. State of the art on ontology alignment. http://knowledgeweb.semanticweb.org/semanticpor tal/deliverables/D2.2.3.pdf

24. Euzenat, J., M. Ehrig and R.G. Castro, 2005. Towards a methodology for evaluating alignment and matching algorithms. http://oaei.inrialpes.fr/doc/oaei-methods.1.pdf

25. Cruz, I.F. and A. Rajendran, 2003. Exploring a new approach to the alignment of ontologies. Proceeding of the Workshop on Semantic Web Technologies for Searching and Retrieving Scientific Data, in Cooperation with the International Semantic Web Conference, Sanibel Island, Florida, pp: 7-12. http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-83/int_1.pdf

26. Huang, J., J. Dang, M.N. Huhns and Y. Shao, 2006. Ontology alignment as a basis for mobile service integration and invocation. J. Pervas. Comput. Commun., 3: 138-158. DOI: 10.1108/17427370710847282

27. Hariri, B.B., H. Abolhassani and A. Khodaei, 2006. A new structural similarity measure for ontology alignment. Proceedings of the International Conference on Semantic Web and Web Services, Las Vegas, USA., pp: 36-42. http://dblp.uni-trier.de/db/conf/swws/swws2006.html#HaririAK06

28. Everingham, M., L. Van~Gool, C.K.I. Williams, J. Winn and A. Zisserman, 2007. The PASCAL visual object classes challenge 2007. http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2007/

29. Ashpole, B., M. Ehrig, J. Euzenat and H. Stuckenschmidt, 2005. FOAM-framework for ontology alignment and mapping; results of the ontology alignment initiative. Proceedings of the Workshop on Integrating Ontologies, Oct. 2005, pp: 72-76. http://sunsite.informatik.rwthaachen.de/Publication s/CEUR-WS/Vol-156/paper11.pdf