Scene of Crime Information System: Playing at St. Andrews

Bogdan Vrusias, Mariam Tariq, Lee Gillam

Department of Computing
University of Surrey, England
{b.vrusias, m.tariq, l.gillam}@surrey.ac.uk

Abstract

This paper discusses the adaptation of the Scene of Crime Information System developed within an EPSRC-funded project, to the collection of data within the ImageCLEF track of the Cross Language Evaluation Forum 2003. The adaptations necessary to participate in this activity are detailed, and initial results are briefly presented.

1. ImageCLEF Collection

ImageCLEF is concerned with the retrieval of images from a specific collection by the captions associated to those images, and is running in relation to an EPSRC-funded project at Sheffield University (Eurovision, GR/R56778/01). The image collection consists of around 28,133 images from the photographic collection provided by St Andrews University Library (Clough et al. 2003). The 28133 images are each referred to and annotated by a single text file, and the full set of annotations are contained within one SGML-based document[1]. Each annotation comprises identifiers to the text file and the image files (DOCNO, SMALL_IMG, LARGE_IMG), the caption of the image (HEADLINE), a set of categories that have been assigned to this image (CATEGORIES), a database record identifier (RECORD_ID) and an unlabelled chunk of text describing the image, denoted below in italics.

<DOC>
<DOCNO>stand03_2093/stand03_27914.txt**</DOCNO>**
<HEADLINE>The Open Championship, St Andrews 1955. Dai Rees and Max Faulkner fishing**.</HEADLINE>**
<TEXT>
<RECORD_ID>GMC-.000007.-.000009.-.000021**</RECORD_ID>**
Rees and Faulkner fishing. Three men in rowing boat tied up at jetty, one holding two fishing rods, one holding oar. July 1955 George Middlemass Cowie Fife, Scotland GMC-7-9-21 mb/
<CATEGORIES>[piers and landing stages],[Fife all views],[rowing boats],[golf - general],[golf - British Open],[rowing],[angling],[battlefields],[fresh water fishing],[fishing vessels],[fishing equipment]**</CATEGORIES>**
<SMALL_IMG>stand03_2093/stand03_27914.jpg**</SMALL_IMG>**
<LARGE_IMG>stand03_2093/stand03_27914_big.jpg**</LARGE_IMG>**
</TEXT>
</DOC>

The information encoded in the XML is intended for use in the retrieval task. By ranked retrieval matching, a set of upto 1000 images is to be retrieved for Task 1, automatic ad hoc retrieval, of the track, and for other purposes in Task 2, interactive image retrieval, of the track.

The above XML fragment refers to the image shown below, of three men in a boat, in Figure 1.

Figure 1: Example Image from the ImageCLEF collection

[1] Although the file was proclaimed to be XML, a number of non-Unicode characters prevented its parsing. It was necessary to replace these with their Hex sequences, ensuring full XML-conformance, to use this collection.

From the above example, it is apparent that some of the categories assigned to the images may not be wholly reliable. While some of the associations are clear:

jetty	piers and landing stages
Fife	Fife all views
rowing (boat)	rowing boats, rowing, fishing vessels
fishing (rods)	angling, fresh water fishing, fishing equipment

others could be associated to information that appears, but is not in the correct context – the combination of "Open Championship" and "St Andrews" being candidates for explaining the golfing categories – while the assignment of a "battlefields" category is less easily obvious.

2. Task 1: Automatic Ad Hoc Retrieval

The automatic ad hoc retrieval task aims at the ranked-retrieval of upto 1000 images from the Eurovision collection. The images are to be retrieved in response to a set of pre-formulated queries. The queries themselves comprise of 50 topics. Each topic has an English query, plus narrative description of the expected result of the query, and the English query has been translated into 5 other languages, French, German, Spanish, Italian and Dutch. Some queries have more than one translation for a given language.

The retrieval results are to be assessed by personnel from the University of Sheffield such that they can be evaluated using the *trec_eval* program with recall and precision metrics. Similar to TREC, the results will subsequently be published.

An example topic encoded in XML[2] is shown below:

```
<top>
<num>Number: 25</num>
<EN-title n="1">Golf course bunkers</EN-title>
<EN-narr>A relevant image will show a picture of a golf course in which a bunker can be clearly identified. The picture must be a photograph or a postcard, but not a drawing, e.g. a plan of the golf course. A bunker is a sandy hollow formed by wearing away of the turf, or nowadays an artificial sand-hole with a built-up face. An example relevant document is [stand03_1714/stand03_7020].</EN-narr>
</top>
<top>
<num>Number: 25</num>
<DE-title n="1">Golfplatz Bunker</DE-title>
</top>
<top>
<num>Number: 25</num>
<FR-title n="1">Bunkers de terrain de golfe</FR-title>
</top>
<top>
<num>Number: 25</num>
<IT-title n="1">Un bunker in un percorso di golf</IT-title>
<IT-title n="2">bunkers in un campo di golf</IT-title>
</top>
<top>
<num>Number: 25</num>
<ES-title n="1">B&#x00FA;nkers en un campo de golf</ES-title>
<ES-title n="2">Pista de golf</ES-title>
</top>
<top>
<num>Number: 25</num>
<NL-title n="1">Bunkers op een golfbaan</NL-title>
</top>
```

The example shown is for Topic 25, for which **Golf course bunkers** has been translated once into each of German, French and Dutch, and twice each for Spanish and Italian. With multiple translations for some languages for the 50 topics, we have the following number of queries for the various languages:

[2] Similar character issues as reported previously were also fixed for this collection.

Spanish	117
English	50
French	51
Italian	103
German	50
Dutch	50
Total	421

These 421 queries are to be made against the 28,133 annotations to retrieve images from the collection.

3. The SoCIS Archetype

The EPSRC-funded Scene of Crime Information System (SoCIS) project was run from October 1999 to March 2003. The aim of the project was to study the link between images and texts within a specialist domain context. A method has been outlined for developing an intelligent content-based image retrieval (CBIR) system, which can store and retrieve images based on the linguistic descriptions of the images. The corpus-based method uses the lexical and semantic properties of specialist texts for extracting key terms and for discovering the ontological organisation of the terms.

A prototype CBIR system was developed in the Java programming language for demonstrating the efficacy of the corpus-based method. The system, which is based on a 3-tier architecture of client, server, and database, can be accessed via a local intranet. SoCIS is an intelligent CBIR system that automatically: (a) labels (and *indexes*) images by keywords as well as relational facts extracted from the descriptions provided by domain experts; (b) extracts physical features of an image; (c) populates a database comprising domain-specific terminology, together with the semantic relationships between terms, starting from a random selection of collateral texts of the domain; and (d) learns to link image and text by using neural networks (Ahmad et al., 2002). SoCIS has integrated modules from (a) System Quirk (Ahmad & Rogers, 2001) - a set of tools for building and managing multilingual term bases with the use of powerful text analysis techniques, and (b) GATE (Cunningham et al., 2002) - a framework and graphical development environment comprising robust NLP tools. The main advantages that SoCIS can be said to have over other text-based and CBIR systems is its ability to extract information from both texts and images, to encode this information for indexing, and to build thesauri, all automatically.

The SoCIS prototype[3] was evaluated using images normally used for the training of Scene of Crime Officers (SoCOs) together with a description provided by the SoCOs as well as other collateral texts like crime scene reports and forensic science research papers and manuals. The question of (inter) indexer-variability, the variances in the output of different indexers for the same image, has been explored in the project (Handy & Ahmad, 2003). This study further reinforced the need for automatic thesauri construction to aid in query expansion (Ahmad et al., 2003a).

4. Adapting SoCIS

SoCIS was specifically targeted at the use of specialist languages – or Languages for Special Purposes (LSP) (Harris, 1988, Ahmad & Rogers, 2001). The system has been built based on the knowledge gathered from Scene of Crime experts, from the testing and evaluation sessions performed with them, and from a domain-specific text corpus. The system had to be adapted to deal with multilinguality as well as structured data from a more general domain for the ImageCLEF collection. SoCIS does not have a translation tool so the translation of the queries from the other languages to English had to be carried out offline as discussed in section 4.1. A parser had to be written to extract the various fields containing textual information (in English) about the images from the provided XML document that could be used for indexing purposes. The indexing module was used to extract single and compound terms from the output of the parser. The main difficulty we encountered (see section 4.2) was the creation of a terminology dictionary and thesaurus related to the general domain, which is needed for the automatic indexing and query expansion modules. We decided to use Wordnet for query expansion purposes but the indexing had to be carried out without using a terminology dictionary to filter out invalid terms. A new relevance ranking mechanism, which is briefly described in section 4.3, was adopted to handle the expanded terms retrieved from Wordnet.

4.1. Handling Multilinguality

The first step necessary was the translation of the various queries to English. Without in-house software, we relied upon translation engines as found on the Internet. Some work was done in an attempt to exploit Google's translation tools for this purpose, however there were difficulties encountered in this. Eventually, Altavista's Babelfish was selected as the principal translation engine (http://babelfish.altavista.com/), however since this system does not translate Dutch, FreeTranslation.com (http://www.freetranslation.com/) was also used.

To translate the queries, Java code was used to wrap definitions of the query syntax used by these sites (with the HTTP POST command being used in both cases). Each query was posted to the site with its requested translation language pair, and the HTML result was retrieved. Using the Java JTidy utility, the resulting HTML was converted to XML (Bray et al, 2000), and XSLT (Clark, 1999) employed to strip out the end result of the translation.

[3] http://www.surrey.ac.uk/socis

The results of translating the various languages for topic number 25 (Golf course bunkers) are shown in the table below:

German	Golf course shelter
French	Bunkers of ground of gulf
Italian (1)	A bunker in a distance of golf
Italian (2)	bunkers in a golf course
Spanish (1)	B??nkers in a golf course
Spanish (2)	Track of golf
Dutch	Bunkers on a wave job

Immediately, certain of these translations will cause problems with the retrieval. The topic identifies the image stand03_1714/stand03_7020 as being relevant. In the run, this was located only for English, Italian (2), and Dutch at ranks 798, 798 and 45 respectively. The quality of returned translation will therefore have a significant impact on the results being returned.

4.2. Synonymy and Morphology

The thesaurus construction module of SoCIS was developed to provide a query expansion facility for the system. There are general-purpose thesauri or lexicons available such as Wordnet[4], which could be used but are inadequate in specialist domains due to a deficiency in specialized terminology. For example, the two key compound terms 'forensic science' and 'crime scene' are not present in Wordnet. The method we developed was based on the analysis of a representative domain-specific text corpus to automatically extract key terms and relationships, which were then used to build the thesaurus (Ahmad et al., 2003a, Tariq et al., 2003). Since the ImageCLEF collection comprised of a wide range of mainly general topics such as buildings, golfers, animals, boats and so on, to apply our method we would have had to construct and analyze a corpus representing most of general knowledge, a clearly difficult and unpractical task. We decided that Wordnet could be a possible resource to use for query expansion since its coverage is based on a general English dictionary.

A program was written to query a Wordnet database to provide a set of synonyms and hyponyms for each of the query terms. In Wordnet, English nouns, verbs, adjectives and adverbs are ordered into synonym sets (synsets). Each synset can be said to contain the words that represent a specific concept. The synsets are then linked to each other based on semantic relations such as antonymy, hyponymy and meronymy. Given a query term, the program returns all the words in the synset that the particular term is an element of, as well as all the hyponyms of each synset element to a specified level in the hierarchy. Initially we planned to go down 2 levels in the hierarchy but ended up using just the synonyms due to system performance issues related to the large number of expanded terms returned, which is discussed in section 5. Taking the query "Boats on Loch Lomond" as an example, the term 'boat' returned 53 expanded words going down one level in the hierarchy. Some synonyms returned were: *travel on water*, *sauceboat*, *gravy boat*; some hyponyms returned included *motorboat*, *mail boat*, *mailboat gondola*, *propel by oars*, *propel by paddles*, *yacht*, and so on. 'Loch' returned one synonym *lough* while 'Lomond' was not present since it is a proper noun. The very common term 'man' had 131 expanded words going down one level and 344 expanded words going down two levels with words such as *private*, *make swollen*, *belly out*, *candy striper*, *Homo erectus*, *clothes horse*, *ridicule with a satire*, and *gentleman*.

Some basic morphological analysis was also carried out for each query term to account for the use of variants such as singular or plural terms as well as the verb or adjective forms. The morphology module uses standard rules (for example if a word ends with 'ss' or 'h' then the plural form is usually derived by adding an 'es') as well as some common exceptions (for example the plural of ~man will be ~men). This was also important for the query expansion part since Wordnet only has singular forms of words as part of the synsets so a plural word used as the query term will return no results.

4.3. Relevance Ranking

Each keyword carried a proportion of its frequency in an annotation divided by the total number of terms allocated to this annotation. The original keyword was then multiplied with weight 1, each expanded term (synonyms) returned by WordNet with weight 0.9, and words containing substrings of the original keywords with weight 0.1. The total ranking was then given by:

$$Rank = \sum \left(\frac{f_{td} \times w_t}{N_d} \right)$$

Where ftd is the term frequency of term t in document d, wt is the weight of a term t as described previously, and Nd is the total number of words in document d.

[4] http://www.cogsci.princeton.edu

5. Performance Issues

The main factor to have an effect on the performance of SoCIS was that the system has been designed for the analysis of free text in specialist domains whereas with the ImageCLEF collection we were dealing with structured texts in a general domain. This resulted in difficulties for SoCIS when indexing the images – the indices produced were relatively unreliable due to the different syntactic structure of the ImageCLEF text when compared to free text, which also affected the ranking. One example here is that the system considered all the category terms given by the ImageCLEF description in the XML document (since they where enclosed in square brackets) as a single compound term. Also due to the fact that we used Wordnet for query expansion, we encountered problems associated with polysemous words as well as different word forms (see the example of *boat* and *man* in section 4.2). Due to the amount of time it was taking to process the expanded queries (some times reaching up to 300 words, see section 4.2) we had to limit the expansion to just synonyms of the original query terms. Even so we had six computers running in parallel to finish the processing, which was taking approximately 8 hours per language.

6. Results and Evaluation

Although the combination of features outlined above would require significant efforts to develop as a usable real-world system (parallelisation and optimisation issues at least), the combination of technologies and techniques presented did enable participation in the ImageCLEF track. A system that in principle would allow a user to query a collection of images that have been annotated in English, using a query in one of six languages has been prototyped from this combination. According to the abstract from the Eurovision project, such a system had not been implemented or researched. Though far from perfect, the evaluation of the results obtained at this stage is important.

Across all languages, the following sets of results were obtained (missing topics and quantities for that topic are given in the third column):

Spanish	105 / 117	32 (3), 33 (1), 34 (1), 36 (1), 39 (2), 43 (3), 47 (1)
English	48 / 50	40, 46
French	47 / 51	7, 17, 25
Italian	91 / 103	13 (2), 17 (1), 27 (3), 29 (2) 31 (1), 39(1), 43 (1), 45 (1),
German	43 / 50	4, 7, 13, 27, 40, 46, 48
Dutch	38 / 50	5, 7, 13, 17, 18, 20, 27, 29, 36, 39, 40, 43
Total	372 / 421	

From a selection of topics, we should evaluate where the exemplar image is ranked and the relevance of the top 10 images retrieved to the query.

	Caption	**Exemplar**
7	Home guard on parade during World War II	stand03_1955/ stand03_24985
14	Boats on Loch Lomond	stand03_1346/ stand03_15600
21	Animals by the photographer Lady Henrietta Gilmour	stand03_1955/ stand03_5603
28	Pictures of golfers in the nineteenth century	stand03_2036/ stand03_7549
35	The mountain Ben Nevis	stand03_1643/ stand03_4692
42	University buildings	stand03_1853/ stand03_21431

	Language and Rank
7	Not found
14	Not found
21	Dutch [884], English [408], Spanish [408, 274, 884], French [408], German [764], Italian [408, 700]
28	Italian [179]
35	French [886], Italian [361]
42	Dutch [971]

For this selection of 6 topics, the exemplar image is only found for English for topic 21. This is an initially disappointing result. We consider, first, the top image being retrieved for each of these topics.

7 (En)	stand03_1749/stand03_22144	Littlehampton. The **Parade**.
14 (En)	stand03_1502/stand03_16737	The Castle, **Loch** an Eilein
21 (En)	stand03_1675/stand03_22740	Engraving of a painting of a Biblical scene, [Noah, family and the Ark at Mount Ararat].
28 (En)	stand03_1714/stand03_7540	Old Tom Morris, **golfer**, St Andrews. (*ca 1900)*
35 (En)	stand03_1851/stand03_7899	Trossachs. Loch Achray, Trossachs Church and Ben An or Binnein (**Ben** A 'an).
42 (En)	stand03_1590/stand03_28349	Samuel Messieux, refugee from Paris and teacher of French at Madras **College** [South Street], St Andrews.

7 (Fr)	No results	
14 (Fr)	stand03_1853/stand03_12134	**Boat** of Garten.
21 (Fr)	stand03_1675/stand03_22740	Engraving of a painting of a Biblical scene, [Noah, family and the Ark at Mount Ararat].
28 (Fr)	stand03_2046/stand03_13818	Kingsbarns. Old Grave Stone, Kingsbarns Churchyard.
35 (Fr)	stand03_1851/stand03_7899	Trossachs. Loch Achray, Trossachs Church and Ben An or Binnein (**Ben** A 'an).
42 (Fr)	stand03_1590/stand03_28349	Samuel Messieux, refugee from Paris and teacher of French at Madras **College** [South Street], St Andrews.

7 (De)	No results	
14 (De)	stand03_1857/stand03_9586	View of **ship** at **sea**.
21 (De)	stand03_1675/stand03_22740	Engraving of a painting of a Biblical scene, [Noah, family and the Ark at Mount Ararat].
28 (De)	stand03_2046/stand03_13818	Kingsbarns. Old Grave Stone, Kingsbarns Churchyard.
35 (De)	stand03_1851/stand03_7899	Trossachs. Loch Achray, Trossachs Church and Ben An or Binnein (**Ben** A 'an).
42 (De)	stand03_2054/stand03_18895	Motherwell. Town Hall.

7 (It)	stand03_1587/stand03_28525	[Walker family?] Untitled portrait of a man.
14 (It)	stand03_1502/stand03_16737	The Castle, **Loch** an Eilein
21 (It)	stand03_1675/stand03_22740	Engraving of a painting of a Biblical scene, [Noah, family and the Ark at Mount Ararat].
28 (It)	stand03_2046/stand03_13818	Kingsbarns. Old Grave Stone, Kingsbarns Churchyard.
35 (It)	stand03_1778/stand03_4502	Launch X.
42 (It)	stand03_2054/stand03_18895	Motherwell. Town Hall.

7 (Es)	stand03_1587/stand03_7524	Man in theatrical costume. [St Andrews ?].
14 (Es)	stand03_1853/stand03_12134	**Boat** of Garten.
21 (Es)	stand03_1675/stand03_22740	Engraving of a painting of a Biblical scene, [Noah, family and the Ark at Mount Ararat].
28 (Es)	stand03_2046/stand03_13818	Kingsbarns. Old Grave Stone, Kingsbarns Churchyard.
35 (Es)	stand03_2092/stand03_14170	Lochgilphead. Crinan Canal at
42 (Es)	stand03_1590/stand03_28349	Samuel Messieux, refugee from Paris and teacher of French at Madras **College** [South Street], St Andrews.

7 (Nl)	stand03_1587/stand03_7524	No results
14 (Nl)	stand03_1502/stand03_16737	The Castle, **Loch** an Eilein
21 (Nl)	stand03_1974/stand03_11773	Brompton Oratory. Altar of Our Lady of Good Counsel.
28 (Nl)	stand03_2046/stand03_13818	Kingsbarns. Old Grave Stone, Kingsbarns Churchyard.
35 (Nl)	stand03_1853/stand03_21295	Fettercairn. Cairn o' Mount and Clatterin' Brig
42 (Nl)	stand03_2054/stand03_18895	Motherwell. Town Hall.

These tables of results show some interesting features. For Topic 7, 3 of the queries returned no results, while those that did have a different first result. For topic 14, 5 of the 6 results refer to just 2 images. For topic 21, all but then Dutch

result refer to the same image. For topic 28, all but the English result refer to the same image, however judging by the caption, the English result is the best. For topic 35, one image is referred to in 3 results. For topic 42, 2 images are equally referred to.

For Topic 14, the top 5 results have been taken once for each language, and the similarity matrix between these results is as follows:

	En	**Fr**	**De**	**Es**	**It**	**Nl**	**Total**
16737	2	1		1	2		**4**
12134	1				1	1	**3**
14211	3				3	2	**3**
22301	4				4	3	**3**
16430	5				5	5	**3**
29031			3	3			**2**
16014			4	4			**2**
12138			5	5			**2**
16009		2					**1**
9586		3					**1**
9587		4					**1**
4618		5					**1**
13150			1				**1**
16833			2				**1**
5702				2			**1**
20573						4	**1**

The top 5 results show degrees of similarity between the English, Italian and Dutch results, with German and Spanish showing similarities, and French showing the most marked behavioural difference. This top 5 have captions as follows:

The Castle, Loch an Eilein
Boat of Garten.
Dunkeld. Loch of Craiglush and Creag nam Mial (Creagnam Hill).
Linlithgow Palace and Loch, from the air.
Bearsden. St Germain's Loch.

It would appear that a number of Lochs, apart from Loch Lomond with any boats on have been discovered in response to this query! Indeed, none of the 16 results above make mention of Lomond.

From this, it is apparent that although similar behaviour is achieved for certain language translations, the end result of retrieval is not correctly weighted. The initial concern that translation would have a significant bearing on retrieval is perhaps now not so relevant as the retrieval itself.

Taking a list of the exemplar images for retrieval, the ranking (where it exists) of that image within the 1000 results for each language was considered. For each language, if the exemplar image was retrieved within the first 1000, this was counted. If it was retrieved within the top 100 results, this was also noted. The following table presents the results obtained.

			High	**Low**	**Ave**	**In top 100**
Nl	20	50	17	971	319.55	7
De	21	50	11	973	309.05	11
En	28	50	8	798	257.96	12
Fr	33	51	11	995	337.52	11
It	42	103	1	967	353.14	14
Es	51	117	7	884	314.52	19

For two queries in Italian, both for Topic 19, the exemplar image was retrieved in first place. This is certainly a result of interest given the analysis of other results in this paper. In the above table, the first column represents the language code, the second the amount of exemplar images retrieved in the 1000 results, the third is the amount of queries, the fourth and fifth show the highest and lowest ranking of the exemplars, with the sixth column showing the average ranking. Column 7 shows the quantity of exemplars occurring in the first 100 retrieved results. This set of results tends to indicate that there is some value to the approach taken here, but how that compares to other approaches remains to be seen.

7. A Note on Text and Image Retrieval

Increasingly, images are being indexed and retrieved by both their visual content and by related texts such as captions that describe the image (Srihari, 1995, Srihari et al., 2000, Paek et al., 1999, Barnard & Forsyth, 2001). Image descriptors extracted directly from image data (colour, texture and shape) tend to capture little of an image's semantic content (Squire et al., 2000, Eakins, 2002) – hence there is a need to extract information about the image content from *collateral* texts (Smeulders et al., 2000, Gillam et al., 2002, Salway & Frehen, 2002, Ahmad et al., 2003a).

8. Future Work

Numerous improvements suggest themselves, for example if the system could be grid-enabled then the different processing modules, as well as instances of the same module, could be run as a service, in parallel, which would significantly improve the processing time. The ranking mechanism needs to be further refined and tuned by carrying out more trial runs. To improve the query expansion one suggestion could be to use part-of-speech information from the query sentence to filter out some of the irrelevant expanded terms returned from Wordnet – for example in the query "Boats on loch Lomond", the term boat is being used in the noun and not the verb form so the synonyms (*propel by oars, propel by paddles)* related to the verb form of *boat* would not be retrieved. Otherwise, an attempt could be made to analyze the British National Corpus[5], which might perhaps yield only the more frequently used term associations.

Since the system deals with image retrieval, we are investigating methods of effectively combining text-based with image-based retrieval techniques. The physical features of an image such as colour, texture, and shape can be extracted and used in combination with the text features. This technique when incorporated into a system that learns how to index, would result in a significant improvement in performance (Ahmad et al., 2003b). We are also investigating the creation of multimedia thesauri, based on Picard's initial work (Picard, 1995). The premise here is that since specialist texts can be said to be a reflection of the ontological commitment of domain experts, specialist images may also reflect some form of ontological commitment on the part of the expert. Also, objects depicted in specialist images often represent the same concepts that are represented by lexical units in texts. The method discussed in Ahmad et al. (2002) could help in establishing the link between an image and text.

9. Acknowledgements

This work was partially funded by the European Union through the Generic Information-based Decision Assistant (GIDA: IST-2000-31123) Projects and the EPSRC through the Scene of Crime Information System (SOCIS: GR/M89041/01) project.

References

Ahmad, K. Vrusias, B., Tariq, M., (2002) "Co-operative neural networks and integrated classification". In *Proceedings of the International Joint Conference on Neural Networks.* Hawaii, USA May 2002. IEEE Press.

Ahmad, K. and Rogers, M., (2001) "Corpus linguistics and terminology extraction". In Wright S.E. and Budin G. (eds.) *Handbook of Terminology Management*, Vol. 2, Amsterdam/Philadelphia: Benjamins, pp. 725-760.

Ahmad, K., Tariq, M., Vrusias, B. and Handy C. (2003a) "Corpus-Based Thesaurus Construction for Image Retrieval in Specialist Domains". In F. Sebastiani (ed.): *Proceedings of the 25th European Conference on Information Retrieval Research*, ECIR-03, Pisa, Italy LNCS-2633. Heidelberg: Springer Verlag. pp 502-510,.

Ahmad, K., Casey, M., Vrusias, B., and Saragiotis, P. (2003b) "Combining Multiple Modes of Information Using Unsupervised Neural Classifiers". In: Windeatt, T. and Roli, F. (eds.), *Proceedings of Multiple Classifier Systems 4th Int. Workshop*, Guildford, UK, June 11-13, 2003, LNCS 2709. Heidelberg: Springer-Verlag, pp. 236-245, 2003b.

Barnard, K., and Forsyth, D. (2001) "Learning the Semantics of Words and Pictures". *International Conference on Computer Vision*, Vol 2, pp 408-415.

Bray, T., Paoli, J., Sperberg-McQueen, C.M., Maler, E. (eds.), (2000). "Extensible Markup Language (XML)", Version 2.0. W3C Recommendation. http://www.w3.org/TR/REC-xml

Clark, J. (ed.), (1999). "XSL Transformations (XSLT)", Version 1.0. W3C Recommendation. http://www.w3.org/TR/xslt

Clough, P., Sanderson, S., Reid, N. (2003). "The Eurovision St Andrews Photographic Collection (ESTA)". http://ir.shef.ac.uk/imageclef/guide.pdf (February 2003)

Cunningham, H. Maynard, D. Bontcheva, K. Tablan, V. (2002) "GATE: A framework and graphical development environment for robust NLP tools and applications". In *Proceedings of the 40th Anniversary Meeting of the Association for Computational Linguistics, 2002*.

[5] http://www.hcu.ox.ac.uk/BNC

Eakins, J.P., (2002) 'Towards intelligent image retrieval'. *Pattern Recognition.* Vol 35, pp 3-14.

Gillam, L., Ahmad, K. and Salway, S. (2002) *Digital Heritage and the use of Terminology*. In Proceedings of 6th International Conference Terminology and Knowledge Engineering (TKE) 2002 ISBN 2-7261-1217-X

Handy, C.J. and Ahmad, K. (2003) "Indexer Variability in Visual Domains". To appear in: *Proceedings of the 13th LSP Conference.*

Harris, Z.S., (1998) "Language and Information". In: Nevin, B. (ed.) *Computational Linguistics*, Vol. 14, No.4, Columbia University Press, New York, pp. 87-90, 1988.

Paek, S., Sable, C. L., Hatzivassiloglou, V., Jaimes, A., Schiffman, B.H., Chang, S-F., and McKeown, K. R. (1999) "Integration of visual and text based approaches for the content labelling and classification of Photographs" *ACM SIGIR'99 Workshop on Multimedia Indexing and Retrieval*, Berkeley, California, USA.

Picard, R.W., (1995) "Towards a Visual Thesaurus". In: Ian Ruthven (ed.) *Springer Verlag Workshops in Computing*, MIRO 95, Glasgow, Scotland.

Salway and Frehen (2002), "Words for Pictures: analysing a corpus of art texts". In *Proceedings of 6th International Conference Terminology and Knowledge Engineering* (TKE) 2002 ISBN 2-7261-1217-X

Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., and Jain, R. (2000) "Content-Based Image Retrieval at the End of the early Years". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 12, IEEE Press, pp. 1349-1380.

Squire, McG.D., Muller, W., Muller, H., Pun, T. (2000) "Content-Based Query of Image databases: Inspirations from Text Retrieval". *Pattern Recognition Letters*, Vol. 21. No. 13-14. Elsevier Science, Netherlands, pp. 1193-1198.

Srihari R.K., (1995) "Use of Collateral Text in Understanding Photos". *Artificial Intelligence Review* (Special Issue on Integrating Language and Vision), Vol. 8, pp. 409-430.

Srihari, R.K. and Zhang, Z. (2000) "Show&Tell: a Semi-Automated Image Annotation System". *IEEE Multimedia*, Vol.7, No. 3, pp. 61-71.

Tariq, M., Manumaisupat, P., Al-Sayed, R. and Ahmad, K., (2003) "Experiments in Ontology Construction from Specialist Texts". To appear in: *Proceedings of EUROLAN Workshop: Ontologies and Information Extraction*, Bucharest, Romania, July 28 -August 08.