

The IAPR Benchmark for Assessing Image Retrieval Performance in Cross Language Evaluation Tasks

Michael Grubinger¹, Clement Leung¹, Paul Clough²

¹ School of Computer Science and Mathematics, Victoria University, Melbourne, Australia
michael.grubinger@research.vu.edu.au, clement@matilda.vu.edu.au

² Department of Information Studies, Sheffield University, Sheffield, UK
p.d.clough@sheffield.ac.uk

Abstract. This paper describes a new image collection created for the CLEF (Cross Language Evaluation Forum) cross-language image retrieval (ImageCLEF) track. This royalty-free collection provides additional evaluation data for ImageCLEF, and in particular provides a new collection for the ad-hoc retrieval task over the existing St. Andrews historic photographic collection. This new image benchmark (referred to as IAPR Benchmark), has grown out of an initiative started by the International Association of Pattern Recognition (IAPR) and consists of approximately 25,000 images from a personal photographic collection provided by the lead author. The construction and composition of this unique image collection are described in this paper, together with information on expected use of this image collection in future ImageCLEF tasks.

1 Introduction

1.1 ImageCLEF and the St. Andrews Photographic Collection

The St. Andrews historic photographic collection (referred to as St. Andrews) has formed the basis of experiments for the past three years in the ad-hoc and interactive task at ImageCLEF [1,2,3], the cross-language image retrieval track run as part of CLEF (the Cross Language Evaluation Forum). The St. Andrews collection contains 28,133 historic photographs provided by St. Andrews University Library [4], which holds one of the most important collections of historic photography in Scotland. Most of the photographs are in black and white and were taken by Scottish photographers or Scottish photographer companies. More information on the St. Andrews collection can be found in [5].

Although the St. Andrews collection has provided a valuable contribution to ImageCLEF, there are a number of limitations to its use which include the following. Firstly, the domain of St. Andrews collection is restricted to mainly photographs specific to Scotland (18,874 images) or the UK (26,728 images) and to life from 100 years ago (most of the pictures were taken before 1940 [5]) therefore affecting both its use and effectiveness as a generic evaluation resource. It is questionable whether

results gained from evaluation on this collection are transferable to other collections therefore limiting conclusions that can be drawn from research with this collection.

Secondly, captions are written solely in English which allows only the possibility of running bilingual tasks where English is the target language; other translation pairs cannot be tested (e.g. translation from Spanish into French). It is not possible to offer different target languages and thereby limiting the translation challenge.

Thirdly, the majority of the images (89 %) in the St. Andrews collection are monochrome or black and white, due to the nature of the collection being a collection of historic photographs. A few are also of poor quality (e.g. too dark, too blurry) and this makes it very difficult to test the usefulness of using visual approaches such as content based image retrieval (CBIR) systems which rely predominately on information regarding colour, texture and shape.

Finally, the biggest problem with the using the St. Andrews collection for comparative evaluation between IR systems is the restriction on copyright that hinders redistribution and further use of researchers outside of the CLEF campaign. The image collection, although free of charge, can just only be legally used having officially registered for ImageCLEF; therefore making it unavailable to other research groups.

1.2 Alternative Image Collections for Evaluation

There is an acknowledged need for photographic collections in the assessment of image retrieval systems performance. Many publications have used the Corel Photo CDs to demonstrate the performance of image retrieval systems. However, this photo collection does not offer any detailed text captions and therefore cannot be used for ImageCLEF anyway. Further, these CDs are quite expensive and copyrighted as well, they do not even exist on the market anymore. Other problems and issues that arise with the use of Corel can be found in [6].

In contrast, the Corbis Database would be a possible alternative for ImageCLEF as the images in this collection are annotated with keywords describing various aspects of the image. Keywords alone, however, are not very precise due to the lack of syntax, and they do not really reflect the real world as studies [7,8] show that when unconstrained from a retrieval task, users tend to create narratives to describe images (and do not use keywords). The Corbis Database, which seems to have very strict copyright regulations too, is therefore not ideal for ImageCLEF either. [9] gives more information on the Corbis Collection and [10] further comments on the inadequacy of keywords.

[11] provides useful links to a number of other image benchmarks and image collections. Unfortunately, most of these benchmark systems furnish images with either too primitive image contents [12,13] or dwell excessively on self similarities with hardly any relationships [14] and are therefore not really suitable for the ImageCLEF tasks either.

2 The IAPR Image Collection

2.1 Collection History

In 2000, the International Association of Pattern Recognition (IAPR) started an effort to create a freely available database with annotated images by presenting a set of recommendations and specifications for such an image benchmark system [15]. Based on these criteria, a first version of a benchmark (consisting of 1000 images, 25 queries, and a collection of performance measures) was set up in 2002 and finally published in 2003 [16]. Due to its origin, it was called *IAPR Benchmark* and will also be referred to as such hereinafter.

Early 2004 saw the refinement, improvement and extension of the IAPR Benchmark to 5,000 images with the help of a new benchmark administration system [17, 18]. By the end of that year, the number of freely available images in the collection had risen to 15,000, followed by first discussions of possible involvement of the IAPR Benchmark at future ImageCLEF tasks.

In 2005, the image annotations have been adapted according to the requirements of ImageCLEF, with a continuously growing number of images in the collection. At the time of writing this paper, the IAPR Benchmark consists of approximately 25,000 images (including corresponding thumbnail versions of the images).

As with the St. Andrews collection, each image is annotated with text and accompanied by a text file describing the contents of the image in (up to three) different languages: English, German, and Spanish. The annotation is encapsulated in SGML and compatible with existing TREC collections thereby enabling a smooth transition from the St. Andrews collection (i.e. existing scripts can still be used).

2.2 Example Images

The collection contains pictures taken in places all over the world and comprises a very good cross-section of still natural images, including lots of sports and action pictures, photos of people, animals, cities, landscapes and many other aspects of life nowadays. Here are a couple of examples of some chosen categories:

sports pictures like the famous tennis player Anna Kournikova serving at The Australian Open 2003, Melbourne (Fig. 1a); Motorcyclist Mika Kallio on his KTM at the 2003 Motorcycle Grand Prix on Phillip Island, Australia (Fig. 1b); a shot from the soccer friendly match Australia vs. Turkey, Telstra Dome, Melbourne, 2004 (Fig. 1c); or an unknown snowboarder jumping, with the mountain Drachenwand in the background, Austria (Fig. 1d);



Fig. 1a. Tennis. **Fig. 1b.** Motorcycling. **Fig. 1c.** Soccer. **Fig. 1d.** Snowboarding

city pictures, including famous sights like Sydney's Opera House and Harbour Bridge (Fig. 2a), Paris' Eiffel Tower (Fig. 2b), partly famous shots like The Strip at night in Las Vegas (Fig. 2c),

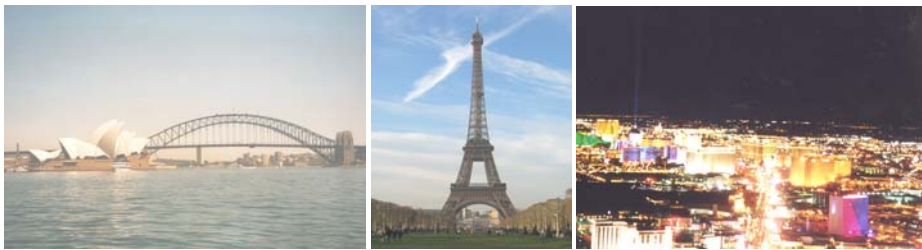


Fig. 2a. Sydney. **Fig. 2b.** Paris. **Fig. 2c.** Las Vegas.

or the statue on the Panacillo in Quito, Ecuador (Fig. 3a); or rather unknown places like a small market in Sahang in Taipei, Taiwan (Fig. 3b) or the skyline of Quito, Ecuador (Fig. 3c);



Fig. 3a. Panacillo. **Fig. 3b.** Taipei. **Fig. 3c.** Quito.

landscape shots, including famous shots like the sunset at Ayers Rock, Australia (Fig. 4a) or the Grand Canyon in Arizona, USA (Fig. 4b); but also not so famous pictures like Lake Mondsee in winter, Austria (Fig. 4c); or this waterfall taken near Stalheimskleiva, Norway (Fig. 4d);

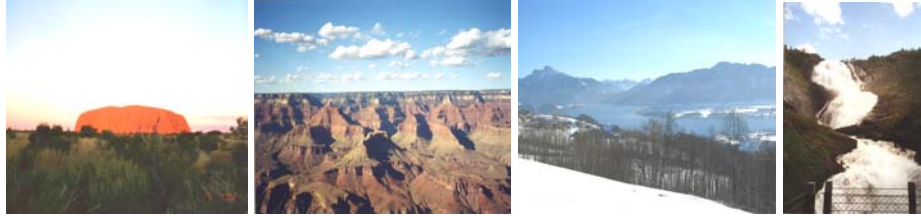


Fig. 4a. Ayers Rock. **Fig. 4b.** Grand Canyon. **Fig. 4c.** Mondsee. **Fig. 4d.** Stalheimskleiva.

animals pictures, including all types of animals like a whale jumping out of the water (Fig. 5a); kangaroos gazing in the grass (Fig. 5b); a cow standing behind a fence (Fig. 5c); or a turtle eating some leaves (Fig. 5d);



Fig. 5a. Whale. **Fig. 5b.** Kangaroos. **Fig. 5c.** Cow. **Fig. 5d.** Turtle.

people shots, like these local children from Arequipa, Peru (Fig. 6a); people at the Love Parade in Berlin, Germany (Fig. 6b); guards posing in Seoul, Korea (Fig. 6c); or a traditional choir group singing in Novosibirsk, Russia (Fig. 6d);



Fig. 6a. Peruvians. **Fig. 6b.** Germans. **Fig. 6c.** Koreans. **Fig. 6d.** Russians.

action shots, like this man pushing his car in winter (Fig. 7a); a girl splashing a cyclist with champagne (Fig. 7b); a boy diving into a swimming pool (Fig. 7c); or a woman drinking out of a glass (Fig. 7d);



Fig. 7a. Pushing.

Fig. 7b. Splashing.

Fig. 7c. Diving.

Fig. 7d. Drinking.

2.3 Text Descriptions

Presently, the text of around 15,000 images is stored in a MySQL database, including a short and a long image description, the place where and the date when the image was taken, and information about the person that took the picture. Most of the annotations are in German, but there are also Spanish and English annotations. Some of them are indexed quite well, others just have an entry for short description or the same entry for short and long description.

Using a benchmark administration system that was presented in [15] and afterwards changed to fulfil the requirements of ImageCLEF (see Fig. 8), these free-text annotations are currently being improved, extended and refined in three different languages: English, German and Spanish.

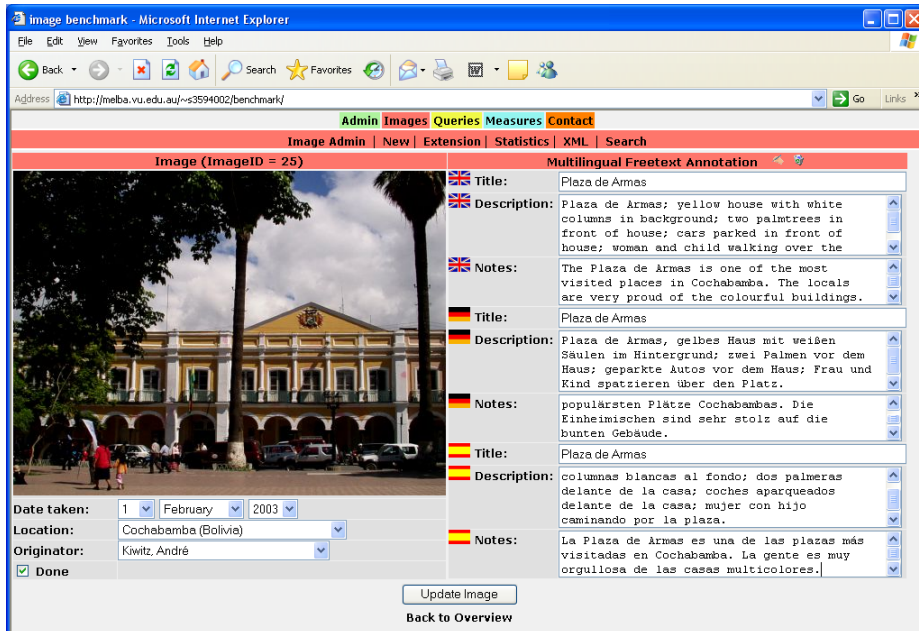


Fig. 8. Screenshot of the new benchmark administration system that is used to update the multilingual free-text annotations.

Each photograph has a caption which consists of the following seven fields: (1) a unique photo number, (2) a title, (3) a free-text description of the image contents, (4) notes for additional information about the photograph, (5) the originator of the photo, (6) the location where the photo was taken e.g. the place/city/town and the country, and (7) the date when the photograph was taken (see Fig. 9).

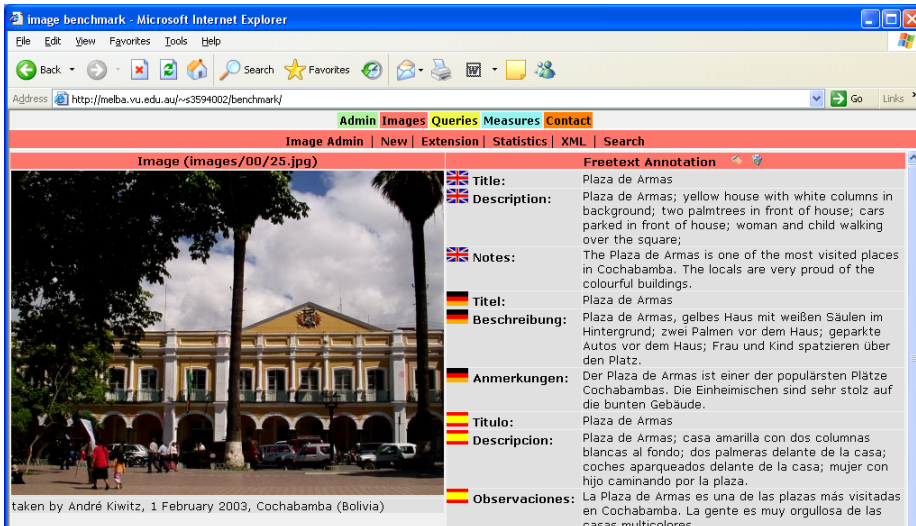


Fig. 9. Benchmark Administration System. Show Image Annotation Page.

With the exception of (1) and (5), all the fields are stored in (up to) three different languages: English, Spanish, and German. This opens up many possibilities for interesting tasks as the organizers of ImageCLEF will be able to generate the text files in any of the three languages for each of the images. Hence, tasks with Spanish or German as source language instead of English will now be possible as well, or even mixed tasks with a random language selection for each of the images (which is very much like what is represented by the Internet as well).

The generated text files contain all captions encapsulated in an SGML format compatible with existing TREC collections, analogue to the format used with the St. Andrews collection in order to keep changes in existing scripts to a minimum. Should the requirements for the format of the text file change in the future, then just the free-text generation script has to be updated and all the information can be generated according to the new requirements very easily. Fig. 10 illustrates the English text file for the image shown in Fig 9.

```

<DOC>
<DOCNO>freetext/00/25.txt</DOCNO>
<LANGUAGE>EN</LANGUAGE>
<HEADLINE>Plaza de Armas</HEADLINE>
<TEXT>
Plaza de Armas; yellow house with white columns in background;
two palm trees in front of house; cars parked in front of
house; woman and child walking over the square.
The Plaza de Armas is one of the most visited places in
Cochabamba. The locals are very proud of the colourful
buildings.
1 February 2003
Andre Kiwitz
Cochabamba, Bolivia
<SMALL_IMG>thumbnails/00/25.jpg</SMALL_IMG>
<LARGE_IMG>images/00/25.jpg</LARGE_IMG>
</TEXT>
</DOC>

```

Fig. 10. Generated image annotation text file in English.

The tags are exactly the same as with the St. Andrews collection [5]. The only difference is the new `<LANGUAGE>` tag used to indicate the annotation language, using the international two-letter abbreviation of the language (e.g. *EN* for English, *DE* for German, *ES* for Spanish, etc.). Fig. 11 shows the same example for German, and Fig. 12 the Spanish equivalent.

```

<DOC>
<DOCNO>freetext/00/25.txt</DOCNO>
<LANGUAGE>DE</LANGUAGE>
<HEADLINE>Plaza de Armas</HEADLINE>
<TEXT>
Plaza de Armas, gelbes Haus mit weißen Säulen im Hintergrund;
zwei Palmen vor dem Haus; geparkte Autos vor dem Haus; Frau
und Kind spazieren über den Platz.
Der Plaza de Armas ist einer der populärsten Plätze
Cochabambas. Die Einheimischen sind sehr stolz auf die bunten
Gebäude.
1. Februar 2003
Andre Kiwitz
Cochabamba, Bolivien
<SMALL_IMG>thumbnails/00/25.jpg</SMALL_IMG>
<LARGE_IMG>images/00/25.jpg</LARGE_IMG>
</TEXT>
</DOC>

```

Fig. 11. Generated image annotation text file in German.

```
<DOC>
<DOCNO>freetext/00/25.txt</DOCNO>
<LANGUAGE>ES</LANGUAGE>
<HEADLINE>Plaza de Armas</HEADLINE>
<TEXT>
Plaza de Armas; casa amarilla con dos columnas blancas al
fondo; dos palmeras delante de la casa; coches aparcados
delante de la casa; mujer con hijo caminando por la plaza.
La Plaza de Armas es una de las plazas más visitadas en
Cochabamba. La gente es muy orgullosa de las casas
multicolores.
1 de febrero de 2003
Andre Kiwitz
Cochabamba, Bolivia
<SMALL_IMG>thumbnails/00/25.jpg</SMALL_IMG>
<LARGE_IMG>images/00/25.jpg</LARGE_IMG>
</TEXT>
</DOC>
```

Fig. 12. Generated image annotation text file in Spanish.

The aforementioned benchmark administration system will not only assist with the generation of these text files, but also be very beneficiary for the topic creation process: Since all the image information is also stored in the database, the estimated target set can automatically be calculated. This helps with the decision whether topics have to be made more specific to reduce the estimated target set or not.

The benchmark administration system will also help observing the complexity of each topic [17]. It is very crucial to provide a good cross section of topic complexity. If topics are too easy, they don't provide a challenge for returning participants. If they are too hard, results are not really meaningful, and especially for new participants it may prove very difficult to obtain good results which might prevent them from presenting their results.

Topics are sent to the participants in a structured format. Once all the topics are stored in the benchmark administration system, these structured text files can easily be generated (without the excessive use of copy and paste!). Should the format of the topic text files change, then just the topic generator has to be changed.

3 Some Statistics and Comparison with St. Andrews

This section provides more information on a number of characteristics of the IAPR collection and in particular a breakdown of image distributions across these characteristics. It also compares the obtained results to those from the St. Andrews collection.

3.1 Image Sizes

There are two different types of dimensions in the IAPR collection, depending on the technology of the capturing device: images that were taken with digital cameras show a 4:3 relation of width and height (96x72 pixels for thumbnails, 480x360 pixels for the larger versions); those taken with a traditional camera (and scanned in later) show a 3:2 relation (92x64 pixels for thumbnails, 480x320 pixels for the larger versions).

As far as storage space is concerned, the thumbnails require between 2 and 5 KB each, the larger versions range from 20 to 200 KB, depending on their contents and colour. All images are stored in jpg format.

Unlike the St. Andrews collection, these images have been modified from the original photos as these would have been too large (1600 x 1200 pixels with digital camera, 1800 x 1200 pixels for scanned photos). The large versions of the photos would range from 200 KB to more than 2 MB, which would have been far too big for the data collection, making the distribution via CDs almost impossible.

3.2 Colour Variation

The majority of images (99.63 %) in the IAPR collection are colour images of very high quality. This is due to the fact that most of these images have been taken in the last five years (many of them with digital cameras), and none of these images is older than 15 years (see also section 3.4). However, there are a couple of images also in black and white (0.37 %). Table 1 compares the colour variation of St. Andrews and IAPR.

Table 1: Comparison of the colour variation of St. Andrews and IAPR

Colour Variation	St. Andrews	IAPR
Colour (%)	11.0	99.6
Black and White / Monochrome (%)	89.0	0.4

In St. Andrews, just one out of 10 photos is not monochrome or black and white. This makes it very hard for the topic creators to include qualifying colour attributes to the topics, which also keeps the topic complexity low. Further, CBIR systems cannot use much colour information for their approach and also struggle with the poor quality of some of the images (too dark, too blurry).

3.3 Variation Across Time

The majority of photographs in the IAPR collection were taken in the new millennium. Fig. 13 shows the distribution of images across the last five years.

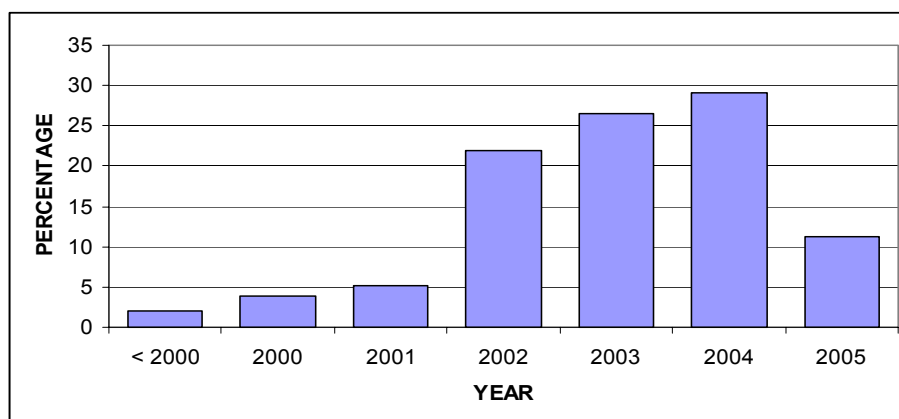


Fig. 13. Variation across time of the IAPR collection for the last five years.

The earliest photo in the collection dates back to 1990, the latest was taken in July 2005 (a range spanning 15 years). The mean year is 2003 (standard deviation is 1.63) and the median is 2003. Table 2 compares these values to the St. Andrews collection:

Table 2: Comparison of the time variation of St. Andrews and IAPR

Variation Across Time	St. Andrews	IAPR
Earliest Photo (year)	1832	1990
Latest Photo (year)	1992	2005
Time Range (years)	160	15
Mean (year)	1920	2003
Median (year)	1931	2003
Standard Deviation (years)	26.2	1.63

Table 2 clearly shows the different nature of both image collections. While the historic St. Andrews collection represents the needs of a certain group of users (like historians), the more modern IAPR collection addresses the needs of a potentially different (and possibly wider) audience. Hence, both collections together provide a good complementary set of evaluation data for image retrieval tasks.

3.4 Variation Across Location

The IAPR Collection is a very diverse collection with pictures taken in more than 40 countries all over the world, including Argentina, Australia, Austria, Bolivia, Brazil, Chile, Colombia, Costa Rica, Croatia, Denmark, Ecuador, England, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Korea, Malaysia, Nicaragua, Norway, Peru, Puerto Rico, Russia, Scotland, Singapore, Spain, Sweden, Switzerland, Taiwan, Trinidad & Tobago, Tunisia, Uruguay, USA, and Venezuela. Fig. 14 shows the variation of the images across these countries.

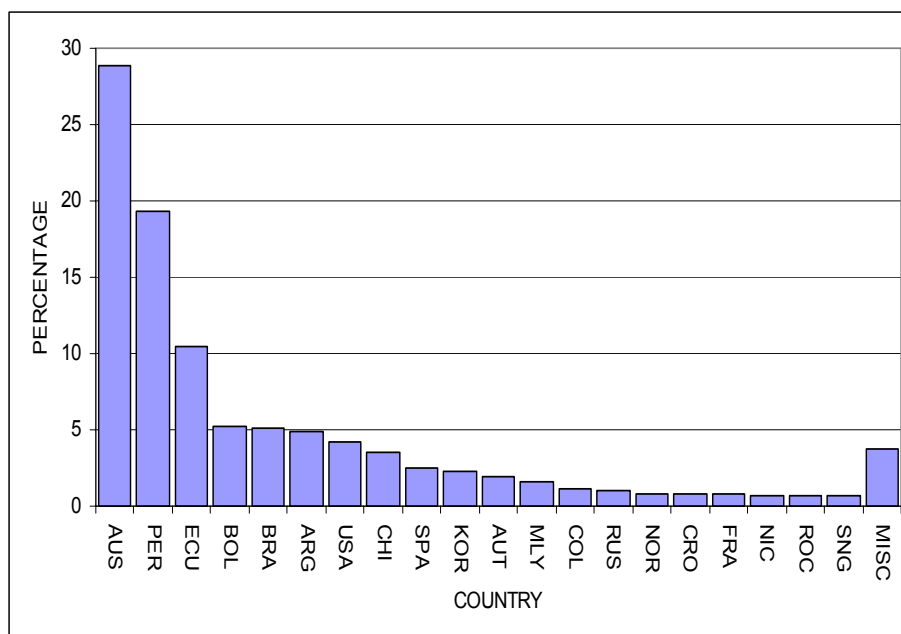


Fig. 14. Variation across countries of the IAPR collection. The countries are represented by their international three letter code.

Most of the images originate from Australia (28.9 %), followed by Peru (19.3 %) and Ecuador (10.5 %). A total of 14 countries contribute to more than 1 % to the collection each, with 24 countries having more than a 100 images. In comparison, 91.4 % of St. Andrews' images were taken in the UK with just only three other countries contributing more than 1 % each. Table 3 outlines this difference in diversity across location of both data collections:

Table 3: Comparison of the variation across location of St. Andrews and IAPR

Variation Across Location	St. Andrews	IAPR
Number of countries in collection	26	40
Percentage of country with most images	91.4	28.9
Countries with more than 1 % of all images	4	14
Countries with more than 100 images	9	24

Table 3 indicates that the St. Andrews collection was very much restricted on just one country, whereas IAPR is spread all over the world. Further, more than 97 % of St. Andrews' images originate from Europe, whereas IAPR comprises a much more international mix of images from South America (50.4 %), Australia (28.9 %), Europe (9.9 %), North America (5.5 %), Asia (5.1 %) and also Africa (0.3 %).

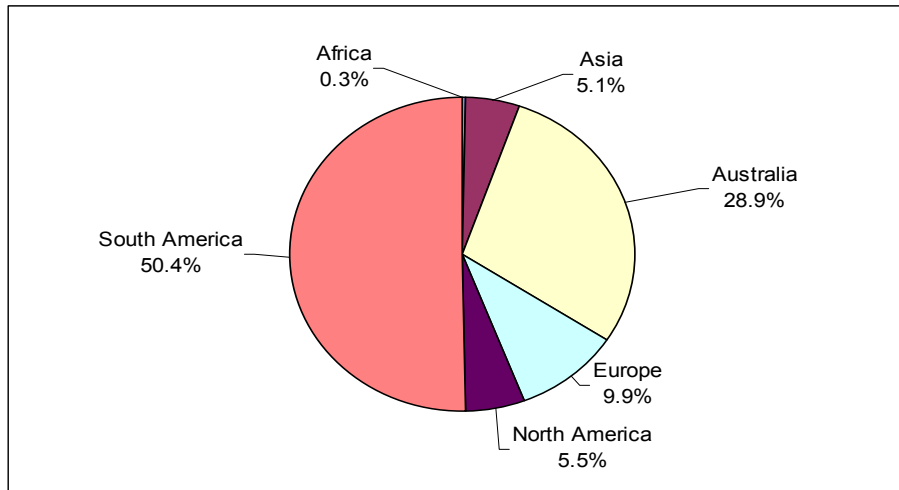


Fig. 15. Variation across continents of the IAPR collection.

4 Integration Into ImageCLEF

This section presents possible topic categories and examples for future ImageCLEF ad-hoc tasks and also outlines possible use for the existing interactive task as well as a possible new task for content based image retrieval only.

4.1 Topics for the existing ad-hoc task

Topics for the bilingual ad-hoc task will be compiled in the same areas as with the St. Andrews collection, including:

Topics modified by location, for example "tourist posing in front of the monument La Mitad del Mundo in Quito, Ecuador" (Fig. 16).

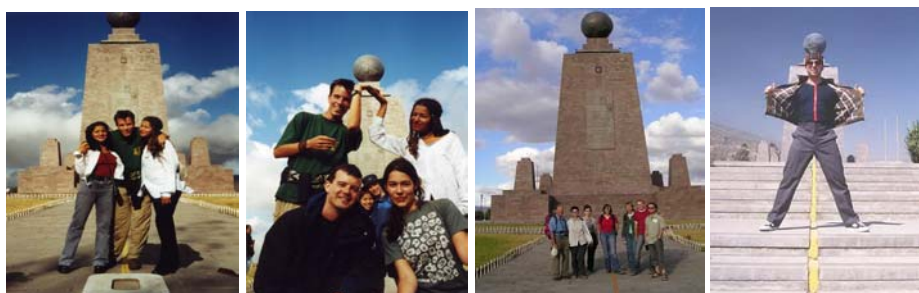


Fig. 16a, 16b, 16c, 16d. Example target images for topic "tourist posing in front of La Mitad del Mundo in Quito, Ecuador".

Topics pertaining to specific events, for example "photos of the Motorcycle Grand Prix 2004, Phillip Island, Australia" (Fig. 17).



Fig. 17a, 17b, 17c. Example target images for topic "Motorcycle Grand Prix 2004, Phillip Island, Australia".

Topics pertaining to known items like "photos of the Iguazu Falls" (Fig. 18):



Fig. 18a, 18b, 18c. Example target images for topic "Iguazu Falls".

General topics like "animal eating", which at the same time tests whether applications can handle general terms – because images might be indexed with *monkey, turtle, dingo, etc.* (Fig. 19).



Fig. 19a, 19b, 19c. Example target images for topic "Animal eating".

More complex topics with subjects and associated attributes (colour, size, etc.), like "a lake with snowy mountain in the background" (Fig. 20):



Fig. 20a, 20b, 20c. Example target images for topic "lake with snowy mountain in the background".

or "cyclist riding mountain-bike downhill" (Fig 21) which does not just ask for a simple subject plus action but also for a manner adverb ("downhill").

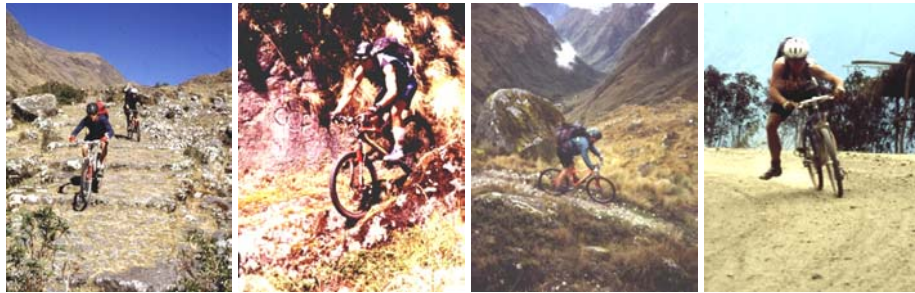


Fig. 21a, 21b, 21c. Example target images for topic "cyclist riding mountain-bike downhill".

Any cyclist riding their bike in flat terrain (or uphill) or people pushing their bikes and so on would not be relevant, or be just partially relevant at best.

4.2 Variations of the existing ad-hoc task

Since annotations are available in three different source languages, there is more scope for additional task scenarios, for example a bilingual task with German (or Spanish) as source language. This would allow the very interesting comparison of search directions (which translation works better, English to Spanish or vice-versa?), and allow more research for translation resources that do not include English (e.g. German to French).

Another possibility would be to provide all three languages randomly (you do not know in advance which image is indexed in what language), a scenario that would best reflect the current situation on the Internet, as a real user does not know which documents are out there in which language either.

4.3 Topics for the existing Interactive Task

The new benchmark could also be used for the interactive task. Also here, the selection process for the target images (topics) would be the same as in the original benchmark. At ImageCLEF 2004 and 2005, some of these challenges were pictures of

- people
- events
- familiar objects (e.g. lighthouses)
- harder objects (e.g. the Tay Bridge rail disaster)
- etc.

Examples for possible target images are given in sections 2.2 and 4.1.

4.4 IAPR and new CBIR Tasks

The results of past ImageCLEF tracks showed that the combination of visual and text based image search shows better results than using the two different approaches individually [19]. Since search approaches at ImageCLEF were still mainly text based in recent years, there is no doubt that test results could be significantly improved if more participants using visual approaches could be attracted. This could be achieved by providing an own task for Content Based Image Retrieval (CBIR), as the topics provided with the bilingual ad-hoc task seem to be very hard for a purely visual approach.

The IAPR collection would be well-suited for such a purely content-based image retrieval task, due to its good quality and colourful and generic nature. Since most of the pictures were taken with a higher resolution than provided in the IAPR collection, the photos could also be distributed with a higher resolution (1200 x 900 pixels), if required.

Further, the IAPR collection could also be used as a complementary benchmark for other CBIR tasks (like the proposed coin identification task) to investigate whether the suggested techniques would also work for images with complex image contents.

5 Advantages of IAPR

The IAPR collection is an excellent addition to the St. Andrews image collection because of the following reasons:

- The domain of the images is very generic, with pictures from all over the world covering a wide cross section of daily life situations.
- All the photos are of good quality as they were mainly captured during the last five years using the newest technology.

- Since more than 99 % of the images are colour images, CBIR approaches based on colour, texture and shape can be tested much better.
- The retrieval of complex image contents is more meaningful as more attributes (colour, etc) can be associated with the nouns where a simple keyword search would not suffice anymore.
- Annotation files can be generated in three different languages. Hence, this collection is not just very well-suited for (cross-language) image retrieval via captions, it opens the door to many new, interesting tasks using other languages than English as source language as well.
- Since the annotation format of the existing annotation file is used, hardly any changes will be necessary for existing scripts.
- The availability of the benchmark administration system will facilitate the topic creation process.
- Finally, and maybe most important, the IAPR collection is free of charge and without any copyright restrictions.

6. Conclusion

This paper presented a new image collection for future ImageCLEF tasks: the IAPR collection, a unique collection of about 25.000 images that are annotated in (up to) three different languages. The domain of the image collection was described, providing a wide range of examples of a broad cross-section of image categories, and the format and creation of the associated text annotations was discussed. The characteristics of the new collections were analysed and compared to the currently used St. Andrews collection, followed by a description of typical topics for existing tasks and ideas for possible new tasks. We have also seen that there are a significant number of advantages of using the IAPR Benchmark.

The IAPR collection is the starting point of a royalty free test collection that will offer researchers world wide the opportunity to experiment and further investigate methods of image retrieval. However, the development of such a benchmark is an incremental and ongoing process and will comprise of further extension, improvement and refinement of this collection. This would, preferably, also include the integration of annotations in other languages too.

Thus, input and participation from researchers in the international research community will be essential to achieving this final goal.

References

1. Clough, P., Sanderson, M.: The CLEF 2003 cross language image retrieval task. In Proceedings of the Cross Language Evaluation Forum (CLEF) 2003, Springer Verlag, (2004).
2. Clough, P., Müller, H., Sanderson, M.: Overview of the CLEF cross-language image retrieval track (ImageCLEF) 2004. In C. Peters, P.D. Clough, G. J. F. Jones, J. Gonzalo, M. Kluck, and B. Magnini, editors, Multilingual Information Access for Text, Speech and

- Images: Result of the fifth CLEF evaluation campaign, Lecture Notes in Computer Science, Springer Verlag, Bath, England (2005).
3. Clough, P., Desealers, T., Grubinger, M., Hersh, W., Jensen, J., Lehmann, T., Müller, H.: The CLEF 2005 Cross Language Image Retrieval Track. In these proceedings. 2005.
 4. Reid, N.: The photographic collections in St. Andrews University Library, Scottish Archives Vol. 5 (1999) 83 – 90.
 5. Clough, P., Sanderson, M., Reid, N.: The Eurovision St Andrews Photographic Collection (ESTA), Image CLEF Report, University of Sheffield, UK (February 2003).
 6. Müller, H., Geissbuhler, A., Marchand-Maillet, S.: The truth about Corel – evaluation in image retrieval. In Proceedings of the International Conference on the Challenge of Image and Video Retrieval (CIVR 2002), London, England (July 2002).
 7. O'Connor, B., O'Connor, M., Abbas, J.: User Reactions as Access Mechanism: An Exploration Based on Captions for Images. Journal of the American Society For Information Science, Vol 50, No 8, (1999).
 8. Jorgensen, C.: The applicability of existing classification systems to image attributes: A selected review. Knowledge Organisation and Change, 5 (1996) 189 – 197.
 9. Janecek, F. P.: Interactive Semantic Fisheye Views for Information Workspaces, PhD Thesis Nr. 2994, Ecole Polytechnique Federale de Lausanne, Switzerland (2004).
 10. Tam, A. M., Leung, C.: Structured Natural-Language Descriptions for Semantic Content Retrieval of Visual Materials. Journal of the American Society for Information Science and Technology, 52/11 (2001) 930 – 937.
 11. Marchand-Maillet, S.: Official Homepage Of VIPER (Visual Information Processing For Enhanced Retrieval), viper.unige.ch, University Of Geneve, Switzerland (October 2002).
 12. Nes, N., Kersten, M.: “The Acoi Algebra: A Query Algebra For Image Retrieval Systems”, University Of Amsterdam, www.cwi.nl/~niels/pub/bncod/paper/paper.html, (1998).
 13. Shapiro, L. G.: Object and Concept Recognition For Content Based Image Retrieval, Ground Truth Database, In Report Nr. IIS-0097329, Department of Computer Science and Engineering, University Of Washington, Seattle, WA, USA (2004).
 14. Gunther, N.J., Beretta, G.B.: A Benchmark For Image Retrieval Using Distributed Systems Over The Internet (BIRDS-I), HP Labs Technical Reports HPL 2000-162, Palo Alto, CA, USA (2000).
 15. Leung, C., Ip, H.: Benchmarking for Content Based Visual Information Search. In Laurini ed. Fourth International Conference on Visual Information Systems (VISUAL'2000) number 1929 in Lecture Notes in Computer Science, Springer Verlag, Lyon, France (November 2000) 442 – 456.
 16. Grubinger, M., Leung, C.: A Benchmark for Performance Calibration in Visual Information Search, In Proceedings of 2003 Conference of Visual Information Systems, Miami, Florida, USA (September 24-26, 2003) 414 – 419.
 17. Grubinger, M., Leung, C.: Incremental Benchmark Development and Administration, In Proceedings of 2004 Conference of Visual Information Systems, San Francisco, California, USA (September 8-10, 2004) 328 – 333.
 18. Over, P., Leung, C., Ip, H., Grubinger, M.: Multimedia Retrieval Benchmarks. Digital Multimedia on Demand, IEEE Multimedia April-June 2004 (2004) 80 - 84.
 19. Müller, H., Clough, P., Hersh, W., Geissbuhler, A.: ImageCLEF 2004 – 2005: Results, Experiences and New Ideas for Image Retrieval Evaluation. Result of the fifth CLEF evaluation campaign, Lecture Notes in Computer Science, Springer Verlag, Bath, England (2005)