

Experiences at ImageCLEF 2010 using CBIR and TBIR mixing information approaches

J. Benavent², X. Benavent², E. de Ves², R. Granados¹, Ana García-Serrano¹

¹ Universidad Nacional de Educación a Distancia, UNED

² Universidad de Valencia

xaro.benavent@uv.es, agarcia@lsi.uned.es

Abstract. The main goal of this paper it is to present our experiments in ImageCLEF 2010 Campaign (Wikipedia retrieval task). This edition we present a different way of using textual and visual information based on the assumption that the textual module better captures the meaning of a topic. So that, the TBIR module works firstly and acts as a filter, and the CBIR system reorder the textual result list. The CBIR system presents three different algorithms: the automatic, the query expansion and a logistic regression relevance feedback algorithm. We have submitted nine textual and eleven mixed runs. Our best run, at the 34th position (25% at the first result list), is a textual run using our own implemented algorithm based on a VSM approach and TF-IDF weights (included in the IDRA tool) and all languages for annotation and for the topics. Our best mixed run (51th position is at 60% first result list) is using the textual list and the logistic regression relevance algorithm at the CBIR module. Most of our runs are above the average of its own modality for the different measures. The new system architecture with the IDRA tool for the textual module and the logistic regression relevance algorithm for the visual module are the right track to maintain in our research lines.

Keywords: Information Retrieval, Textual-based Retrieval, Content-Based Image Retrieval, Relevance feedback, Merge Results Lists, Fusion, Indexing.

Categories and subject descriptors

H.3 [Information Storage and Retrieval]: H.3.1 Content Analysis and Indexing; H.3.2 Information Storage; H.3.3 Information Search and Retrieval; H.3.4 Systems and Software; H.3.7 Digital libraries. **H.2 [Database Management]:** H.2.5 Heterogeneous Databases; **E.2 [Data Storage Representations].**

1 Introduction

The UNED-UV is a research group formed by researchers from two different universities in Spain, the Universidad Nacional de Educación a Distancia (UNED) and the Valencia University (UV). This research group is working together [1] [2] since ImageCLEF08 edition.

The main goal of this paper is to present our experiments in ImageCLEF 2010 Campaign (Wikipedia retrieval task) [3]. This ImageCLEF edition our group presents a different way of working using the information of the Content Based Image Retrieval (CBIR) system and the information of the Textual Based Image Retrieval (TBIR) system. The global system is based on the assumption that the conceptual meaning of a topic is initially better captured by the text module itself than by the visual module. Therefore, the TBIR system works firstly over the whole database working as a filter, and then the CBIR system reorders the filtered textual result list. In this way, the CBIR system acts also as a merging module.

The TBIR subsystem includes the UNED own implemented tool IDRA (InDexing and Retrieving Automatically) [4] that includes several functionalities, including an algorithm based on the Vector Space Model (VSM) approach using TF-IDF weighted vectors. The CBIR subsystem includes three different algorithms: the automatic, the query expansion and UV own logistic regression relevance feedback algorithm [5].

A more detailed presentation of the system, the submitted experiments, and the obtained results is included in the following sections.

2 System Description

The global system (shown at Fig. 1) includes three main subsystems: the TBIR, the CBIR and the merging module. The TBIR subsystem uses the UNED own implemented tool IDRA [5], in charge of indexing and retrieving textual annotations from images. The Valencia University CBIR system implements for this ImageCLEF edition three different algorithms: the automatic, the query expansion and the relevance feedback algorithm based on logistic regression [5]. The TBIR subsystem acts firstly over the whole images of the database, acting as a filter to the CBIR system selecting the relevant images for a certain query. In a second step, the CBIR system works over the set of filtered images reordering this list taking into account the visual information of the image. The CBIR system generates different visual result lists depending on the number of query images (for the automatic and the query expansion algorithm). These lists are merged by the merging module by an OWA operator [6].

2.1 Text-based Index and Retrieval

This module is in charge of the textual image retrieval using the metadata supplied for the images in the collection. IDRA tool [4] extracts, selects, preprocesses and indexes the metadata information, for later search and retrieve the most relevant images for the queries. After this process, a ranked results list is obtained for each textual experiment.

The textual retrieval task architecture can be seen in the Fig. 1. Each one of the components takes care of a specific task. These tasks will be sequentially executed:

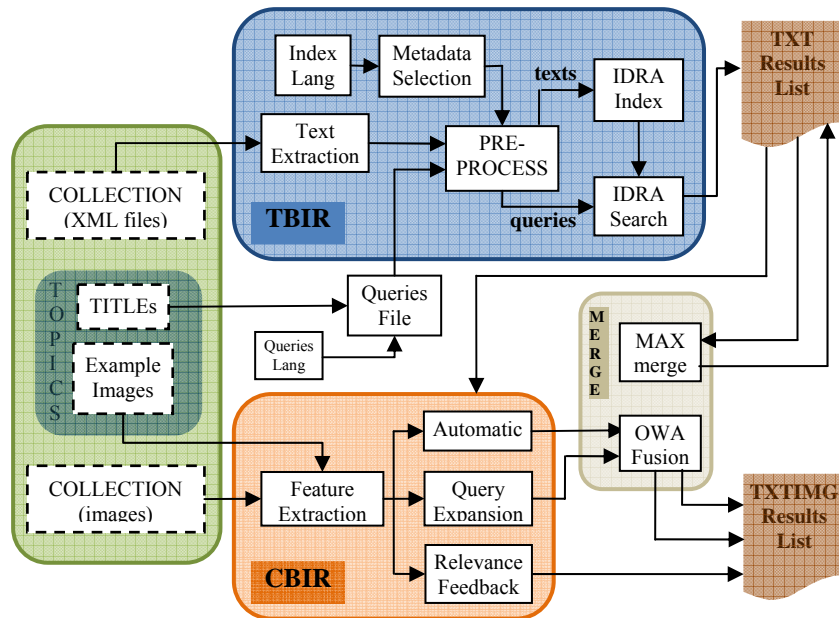


Fig. 1. System overview.

Text Extraction. Extracts the text from the files which contains the associated metadata. It uses the JDOM Java API to identify the content of each of the tags of the XML files.

Preprocess. This component process the text in two ways: 1) special characters deletion: characters with no statistical meaning, like punctuation marks, are eliminated; and 2) stopwords detection: exclusion of semantic empty words from specifics lists for each language. When processing multilingual text, a manually join of these lists is used.

Metadata Selection. With this component the system selects the text we want to index, depending on the chosen language “Index Lang” (EN, FR, DE or ALL). Therefore, 4 different indexations will be generated: one multilingual and three monolingual.

In the case of the monolingual indexations the selected text for the chosen language $L = \{EN, FR, DE\}$, from images metadata files, will be: 1) <name> whenever there is specific metadata for language L, or when there is not for any language; 2) <description> and <caption> whenever there is specific metadata for L; 3) <comment> when the text in this tag is not contained in <description> or <caption> and therefore will add new information. This time we did not use the text from the corresponding Wikipedia articles indicated in the <caption> attribute “article”.

When carrying out the multilingual indexation (ALL), the selected text will be the concatenation of the corresponding text for each of the three languages (EN+FR+DE), in the same way as explained for the monolingual cases.

Queries File. 4 different queries files are constructed for the experiments: one for each language (EN, FR, DE), and another for the multilingual case, which is indicated in “Queries Lang”. The strategy to select the text for each query is just to extract the information in the <TITLE> tag for the chosen language, and the concatenation of the three languages for the multilingual experiment.

IDRA Index. This component indexes the selected text associated with each image. The indexation is based on the VSM approach using TF-IDF (term frequency – inverse document frequency) weighted vectors. This approach consists in calculating the weights vectors for each one of the images selected texts. Each vector is compounded by the TF-IDF weights values of the different words in the collection. TF-IDF weight is a statistical measure used to evaluate how important a word is to a text in a concrete collection, and is calculated as shown in (1).

$$TF - IDF = t_{i,j} * \log_2 \left(\frac{N}{n_i} \right) \quad \begin{array}{l} t_{j,j}: \text{number of occurrences of the word } t_j \text{ in caption text } T_i. \\ N: \text{total number of images captions in the collection.} \\ n_i: \text{number of captions in which appears the word } t_i. \end{array} \quad (1)$$

All weights values for each vector will be then normalized using the Euclidean distance. Therefore, for each one of the words appearing in the collection, the IDRA Index process updates and stores the following values: n_i , $t_{i,j}$, N (described in (1)), T_i : unique identifier of the image, idf_j : inverse document frequency ($\log_2(N/n_i)$) in T_i , E_i : Euclidean distance used to normalize, and $w_{j,i}$: weight of word t_j in T_i .

IDRA Search. Is in charge of launching the queries against a concrete indexation for the experiment, and it obtains the corresponding “TXT Results List”. For each one of the queries, IDRA calculates its corresponding weights vector in the same way as in indexation. Then, the similarity between the query and an image text will depend on the proximity of their associated vectors calculated by the cosine measure:

$$sim(T_i, q) = \cos(\vec{O}) = \frac{\sum w_{j,i} * w_{j,i}}{\sqrt{\sum w_{j,i} * w_{j,i}} * \sqrt{\sum w_{q,i} * w_{q,i}}} \quad (2)$$

This similarity value will be calculated between the query and all the images metadata indexed. Images are ranked in descending order in the “TXT Results List”.

2.2 Content-Based Information and Visual Retrieval

The VISION-Team at the Computer Science Department of the University of Valencia has its own CBIR system, and that has also been used in previous ImageCLEF editions (Photo-retrieval task in 2008 and 2009 [1] [2]). The low-level features of the CBIR system have been adapted for the images of the new image collection (WikipediaMM 2010) taking into account the results of the previous editions.

As in most CBIR systems, a feature vector represents each image. The first step at the Visual Retrieval system is extracting these features for all the images on the database as for each of the cluster query topic images for each question. Instead of using the low-level features provided by the organization, we have used our own

features. We use different low-level features describing color and texture to build a vector of features. The number of low-level features has been increased from the 114 components at ImageCLEF09 up to 296 components at the current edition. This increment is mainly due to the use of local HS histogram (10x3 bins) instead of local H histograms (10 bins) descriptors in previous editions.

- **Color information:** Color information has been extracted calculating both local and global histograms of the images using 10x3 bins on the HS color system. Local histograms have been calculated dividing the images in four fragments of the same size. A bidimensional HS histogram with 10x3 bins is computed for each patch. Therefore, a feature vector of 30 components for the global histogram, and 192 components for the local histograms represent the color information of the image.
- **Texture information:** Two types of texture features are computed: The granulometric distribution function, using the coefficients that result of fitting the distribution function with a B-spline basis. And, the Spatial Size Distribution. We have used two different versions of it by using as the structuring elements for the morphological operation that get size both a horizontal and a vertical segment [1].

At this edition, the vision team has focus his work in testing three different visual algorithms applied to the results retrieved by the text module: the automatic, the relevance feedback and the query expansion. We assume that the conceptual meaning of a question is better captured by the text module than by a visual module when they work individually. Therefore, the task of the visual module is to re-order the textual result list taking into account the information of the query images given at each topic.

Automatic algorithm. This is the typical algorithm in a CBIR system. The first step is to calculate the feature vector that describes each image of the database as it has been explained at the previous paragraph. The second step is to calculate the similarity measurement between the feature vectors of each image on the database and the N query images. The distance metric applied in our experiments is the Mahalanobis distance that gives better results than the Euclidean one ([1]). The Mahalanobis distance gives better results than the Euclidean due to the fact that this measure takes into account the correlations of the data set and is scale-invariant being this characteristic very useful because of the broad differences between the different low-level feature values. The Mahalanobis distance needs to pre-calculate the covariance matrix of the sample data. Since, the size of the database is too huge we have chosen a different approach: a covariance matrix is computed for each textual result list given for each topic. Thus, we have managed to cope with the problem of computing the metric for the Mahalanobis distance in a large database.

As we have N query images, we will obtain N visual result lists, one for each query image in the topic. These N result lists are passed to the merging module to fuse them in one result list.

Query expansion algorithm. The query expansion algorithm works in the same way that the automatic algorithm, being the only difference that this algorithm expands the N query images to a wider set of images M. Thus, the M query images set is

composed of the N images given by the topic and the N' expanded images being $M=N+N'$. The N' images set are the 3 first images of the textual result list. The M result lists are passed to the merging module.

Relevance feedback algorithm based on logistic regression. This algorithm works differently to the two previous ones. Therefore, we will explain the concept of relevance feedback and the adjustments made to get a good performance of the algorithm for the proposed tasks [5]. Relevance feedback is a term used to describe the actions performed by a user to interactively improve the results of a query by reformulating it. An initial query formulated by a user may not fully capture his/her wishes. Users then typically change the query manually and re-execute the search until they are satisfied. By using relevance feedback, the system learns a new query that better captures the user's need for information. The user enters his/her preferences at each iteration through the selection of relevant and non-relevant images.

We will explain the way the logistic regression relevance feedback algorithm works. Let us consider the (random) variable Y giving the user evaluation where $Y=1$ means that the image is positively evaluated and $Y=0$ means a negative evaluation. Each image in the database has been previously described by using low-level features in such a way that the j -th image has the k -dimensional feature vector x_j associated. Our data will consist of (x_j, y_j) , with $j=1, \dots, n$, where n is the total number of images, x_j is the feature vector and y_j the user evaluation (1 =positive and 0 =negative). The image feature vector x is known for any image and we intend to predict the associated value of Y . In this work, we have used a logistic regression where $P(Y=1|x)$ i.e. the probability that $Y=1$ (the user evaluates the image positively) given the feature vector x , is related with the systematic part of the model (a linear combination of the feature vector) by means of the logit function. For a binary response variable Y and p explanatory variables X_1, \dots, X_p , the model for $\pi(x)=P(Y=1|x)$ at values $x=(x_1, \dots, x_p)$ of predictors is $\text{logit}[\pi(x)]=\alpha+\beta_1x_1+\dots+\beta_px_p$, where $\text{logit}[\pi(x)]=\ln(\pi(x)/(1-\pi(x)))$. The model parameters are obtained by maximizing the likelihood function given by:

$$l(\beta) = \prod_{i=1}^n \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i} \quad (3)$$

The maximum likelihood estimators (MLE) of the parameter vector β are calculated by using an iterative method.

We have a major difficulty when having to adjust a global regression model in which we take the whole set of variables into account, because the number of selected images (the number of positive plus negative images) is typically smaller than the number of characteristics. In this case, the regression model adjusted has as many parameters as the number of data and many relevant variables could be not considered. In order to solve this problem, our proposal is to adjust different smaller regression models: each model considers only a subset of variables consisting of semantically related characteristics of the image. Consequently, each sub-model will associate a different relevance probability to a given image x , and we face the question of how to combine them in order to rank the database according to the user's preferences. This problem has been solved by means of an ordered averaged weighted operator (OWA) [6].

In our case, we have adapted the manual relevance feedback to an automatic performance. The examples and the counter-examples (positive and negative images) are automatically selected for each topic. The examples are the query images of the topic plus N images taken from the first positions of the textual result list. The counter-examples are the M latest positions of the textual result list. The relevance feedback algorithm is executed once.

2.4 Merging Algorithms

Two merging algorithms are used in different steps with different purposes.

OWA Fusion. In the modality for textual and visual retrieval the approach that follows this edition is based on the assumption that the conceptual meaning of a topic is initially better captured by the text module itself than by the visual module. Thus, the textual module works as a filter for the visual module, and the work of the visual module is to re-order the textual results list. In this way, there has not been used an explicit fusion algorithm to merge the textual result list and the visual result list. However, the visual module generates N result visual lists depending on the number of query images for the automatic and query expansion algorithms. These N lists are merged in one result final list by using the Mathematical aggregation operators OWA [6]. The OWA transform a finite number of inputs into a single output and play an important role in image retrieval. With the OWA operator no weight is associated with any particular input; instead, the relative magnitude of the input decides which weight corresponds to each input. In our application, the inputs are similarity distances to each of the N query images and this property is very interesting because we do not know, a priori, which image of the N images will provide us with the best information. The aggregation weights used for these experiments are the weights which correspond to the maximum, that is an OR operator.

MAXmerge. This algorithm is used to fuse together different results lists in order to carry out some experiments related to multilingualism (UNED-UV8, UNED-UV9 described in next section). MAXmerge algorithm is included in IDRA tool and consists on, for each query, to select the results from the different lists which have a higher relevance/similarity value for the corresponding query, independently of the list the results appears in.

3 Experiments (submitted runs)

We have participated in two modalities: textual and mixed retrieval (visual and textual). Finally, 20 runs were submitted (9 textual, 11 mixed). A schematic description of these runs is shown in Table 1.

For textual modality, we present 9 runs. As it is explained in previous sections, 4 different indexations and 4 queries files were generated. From all possible combinations, we were interested in evaluate experiments with the 4 queries files against the multilingual indexation, obtaining 4 runs: UNED-UV1 (with multilingual

queries), UNED-UV2 (with English queries), UNED-UV4 (with French queries), and UNED-UV6 (with Dutch queries). UNED-UV3, UNED-UV5 and UNED-UV7 correspond to monolingual experiments in which the language for the indexation is the same of the queries: English, French or Dutch, respectively. Finally, 2 more textual runs were submitted using the MAXmerge fusion algorithm: UNED-UV8 merging results lists from UNED-UV2, UNED-UV4 and UNED-UV6; and UNED-UV9 merging results from UNED-UV3, UNED-UV5 and UNED-UV7.

In the mixed modality, the textual module has passed through the visual module four different kinds of its basic textual algorithms corresponding to the UNED-UV1, UNED-UV2, UNED-UV3 and UNED-UV9 runs. The visual module has applied its three different algorithms (automatic, relevance feedback and query expansion) to the textual result lists in order to test the performance of these three algorithms over the different kind of text algorithms retrieval. For the [UNED-UV1] basic line the automatic, relevance feedback and query expansion algorithms have been applied getting the three corresponding runs [UNED-UV10], [UNED-UV11] and [UNED-UV12]. Following the same structure applying the three different visual algorithms to the [UNED-UV2] run the [UNED-UV13], [UNED-UV14] and [UNED-UV15] runs are obtained. From the [UNED-UV3] run the [UNED-UV16], [UNED-UV17] and [UNED-UV18]; and, from the [UNED-UV9] the [UNED-UV19], [UNED-UV20] and [UNED-UV21]. The last one was out of the maximum runs submitted.

Table 1. Submitted textual and mixed experiments.

Run	Mod	CBIR		TBIR	
		Algor	Algorithm	Annotation language	Topic language
UNED-UV1	Text	-	VSM	EN+FR+DE	EN+FR+DE
UNED-UV2	Text	-	VSM	EN+FR+DE	EN
UNED-UV3	Text	-	VSM	EN	EN
UNED-UV4	Text	-	VSM	EN+FR+DE	FR
UNED-UV5	Text	-	VSM	FR	FR
UNED-UV6	Text	-	VSM	EN+FR+DE	DE
UNED-UV7	Text	-	VSM	DE	DE
UNED-UV8	Text	-	VSM (EN+FR+DE) + MAXmerge	EN+FR+DE	EN+FR+DE
UNED-UV9	Text	-	VSM (EN FR DE) + MAXmerge	EN+FR+DE	EN+FR+DE
UNED-UV10	Mixed	AUTO	[UNED-UV1]	EN+FR+DE	EN+FR+DE
UNED-UV11	Mixed	FB	[UNED-UV1]	EN+FR+DE	EN+FR+DE
UNED-UV12	Mixed	QE	[UNED-UV1]	EN+FR+DE	EN+FR+DE
UNED-UV13	Mixed	AUTO	[UNED-UV2]	EN+FR+DE	EN
UNED-UV14	Mixed	FB	[UNED-UV2]	EN+FR+DE	EN
UNED-UV15	Mixed	QE	[UNED-UV2]	EN+FR+DE	EN
UNED-UV16	Mixed	AUTO	[UNED-UV3]	EN	EN
UNED-UV17	Mixed	FB	[UNED-UV3]	EN	EN
UNED-UV18	Mixed	QE	[UNED-UV3]	EN	EN
UNED-UV19	Mixed	AUTO	[UNED-UV9]	EN+FR+DE	EN+FR+DE
UNED-UV20	Mixed	FB	[UNED-UV9]	EN+FR+DE	EN+FR+DE

4 Results

After the evaluation by the task organizers, our results for each of the submitted experiments are presented in Table 2. The table shows that our two best results are for the textual runs UNED-UV1 and UNED-UV9 (at the 34th and 40th position of the global result list, this is at the 25% first results). For the mixed modality, the best result is the UNED-UV11 at the 51th position (at the 60% first results). It is worth pointing out that the ranking position is computed by using the MAP measure (with a maximum MAP value of 0.1927 for our best run and a minimum MAP value of 0.1502 for our worst one). It can also be observed at Table 2 that most of our runs are above the average for each own modality (textual and mixed runs). These above results are marked in bold at the table.

Table 2. Results for the submitted experiments (The results in bold are above the average for the modality).

Po	Run	Mode	MAP	P@10	P@20	R-prec.	Bpref	NDCG
34	UNED-UV1	Text	0.1927	0.3914	0.3564	0.2663	0.2282	0.4092
40	UNED-UV9	Text	0.1865	0.4200	0.3636	0.2638	0.2253	0.4012
51	UNED-UV11	Mixed	0.1792	0.3914	0.3629	0.2514	0.2175	0.3887
52	UNED-UV8	Text	0.1790	0.3914	0.3350	0.2533	0.2150	0.4006
59	UNED-UV20	Mixed	0.1717	0.4071	0.3571	0.2499	0.2133	0.3803
61	UNED-UV2	Text	0.1627	0.3657	0.3293	0.2340	0.2002	0.3582
68	UNED-UV12	Mixed	0.1525	0.3943	0.3621	0.2236	0.1939	0.3341
69	UNED-UV10	Mixed	0.1502	0.3971	0.3607	0.2204	0.1920	0.3318
70	UNED-UV14	Mixed	0.1498	0.3543	0.3250	0.2203	0.1902	0.3387
72	UNED-UV19	Mixed	0.1427	0.4171	0.3671	0.2166	0.1872	0.3219
76	UNED-UV3	Text	0.1370	0.3871	0.3336	0.2146	0.1787	0.3168
77	UNED-UV15	Mixed	0.1286	0.3829	0.3386	0.1947	0.1687	0.2935
78	UNED-UV17	Mixed	0.1285	0.3614	0.3379	0.2047	0.1723	0.3049
79	UNED-UV13	Mixed	0.1261	0.3857	0.3307	0.1879	0.1650	0.2909
83	UNED-UV16	Mixed	0.1089	0.4043	0.3357	0.1728	0.1491	0.2588
84	UNED-UV18	Mixed	0.1077	0.3886	0.3307	0.1729	0.1492	0.2571
88	UNED-UV6	Text	0.0936	0.2671	0.2314	0.1312	0.1151	0.1885
89	UNED-UV4	Text	0.0920	0.2829	0.2536	0.1492	0.1301	0.2128
97	UNED-UV5	Text	0.0661	0.2943	0.2650	0.1156	0.1017	0.1703
102	UNED-UV7	Text	0.0603	0.2586	0.2221	0.0994	0.0851	0.1378
Average		Text	0,1579	0,3961	0,3519	0,2277	0,1992	0,3622
Best (pos 12)		Text	0,2361	0,4871	0,4393	0,3077	0,2694	0,5217
Average		Mixed	0,1387	0,3701	0,3293	0,1982	0,1759	0,3319
Best (pos 1)		Mixed	0,2630	0,6110	0,5410	0,3289	0,2970	0,5360

With textual experiments this campaign we aimed to analyze multilingual issues. Comparing UNED-UV1 results with UNED-UV2, UNED-UV4 and UNED-UV6 ones, we can observe that best retrieval with the multilingual indexation is performed when we use the queries file constructed with the concatenation of all languages (MAP=0.1927). Launching English queries obtains better results (0.1627) than French

(0.0920) or Dutch (0.09936), surely due to the metadata information for that language. Analyzing results for UNED-UV8 and UNED-UV9 runs, we observe that both of them obtain a good performance (only UNED-UV1 obtains higher MAP than them). Slightly higher results are for UNED-UV9 ($0.1865 > 0.1790$), so it is early to conclude when merging results from different languages, if it is better to launch the queries against monolingual indexations than against multilingual. At this moment, the effort in preprocessing has to be taken into account to decide.

The process to analyze our mixed modality results is by the comparison of the basic textual algorithms (UNED-UV1, UNED-UV2, UNED-UV3 and UNED-UV9) and their corresponding mixed runs (UNED-UV10-12 for the UNED-UV1, UNED-UV13-15 for UNED-UV2, and so on). We have improved the precision values at 10 and at 20 with the mixed runs, i.e. the UNED-UV11-Feedback (Prec@10 0.3914 and Prec@20 0.3629) improves its basic textual run UNED-UV1 (Prec@10 0.3914 and Prec@20 0.3564). The same improvement can be observed for the other mixed runs being compared with their corresponding textual runs. This result points out that visual algorithms can improve the textual result lists by back forward to the end of the list non relevant images retrieved by the textual module. However, MAP values are still lower than their corresponding textual runs. This could be due to the fact that more query images would be needed to get better results for higher precision measures (P@30, P@40 and so on), improving in that way the medium of the precision values (MAP).

5 Concluding Remarks and Future Work

Our best result is for the textual modality and it is at the position 34th, at the first 25% of the best results of the contest; and, our best result for the mixed modality is at the 51th position, at the first 60% of the global contest. Most of our runs in ImageCLEF10 are above the average for its own modality. These results mean that our main algorithms for textual and visual modules have got good marks, and they can be tuned to improve the current results.

Regarding multilinguality, the multilingual run it is our best (multilingual query launched to the multilingual index). It defeats the runs using monolingual queries (also on multilingual index). When using monolingual indexes and merging the results lists according to the query, only a slightly difference it is obtained (MAP value $0.1927 > 0.1865$ for UNED_UV9). It is early to conclude about, but the effort in preprocessing has to be taken into account to it.

The best result for mixed runs has been obtained with the logistic regression relevance feedback algorithm (UNED-UV11 at position 51), followed by the query expansion and the automatic one. Our new algorithms (logistic regression relevance feedback and query expansion) have markedly improved the results in comparison with the automatic algorithm used in previous editions. It is also important to notice that the best results of the contest are also achieved with a feedback algorithm. This reinforces our idea that the feedback algorithms are the right track to maintain in our future research lines.

Acknowledgments. This work has been partially supported by projects TIN2007-67407-C03-03, TIN2007-67587 and TEC2009-12980 from Spanish government.

References

1. Ana García-Serrano, Xaro Benavent, Rubén Granados, José Miguel Goñi-Menoyo. Some results using different approaches to merge visual and text-based features in CLEF'08 photo collection. Lecture Notes in Computer Science, Evaluating Systems for Multilingual and Multimodal Information Access. Vol.: 5706/2009 Págs, 568-571. ISSN: 0302-9743.
2. R. Granados, X. Benavent, R. Agerri, A. García-Serrano, J.M. Goñi, J. Gomar, E. de Ves, J. Domingo, G. Ayala. MIRACLE (FI) at ImageCLEFphoto 2009. Cross-Language Evaluation Forum CLEF 2009. Working Notes for the CLEF. Corfu (Grecia), September 2009.
3. Adrian Popescu, Theodora Tsirikika, and Jana Kludas. Overview of the Wikipedia Retrieval task at ImageCLEF 2010. In the Working Notes of CLEF 2010, Padova, Italy, 2010.
4. Rubén Granados Muñoz, Ana García Serrano, José M. Goñi Menoyo. La herramienta IDRA (Indexing and Retrieving Automatically). Procesamiento del Lenguaje Natural, nº 43, Septiembre de 2009. XXV Conferencia de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN'09). San Sebastián, 2009.
5. Leon, T., Zuccarello, P., Ayala, G., de Ves, E., Domingo, J.: Applying logistic regression to relevance feedback in image retrieval systems, Pattern Recognition, vol. 40, pp. 2621--2632. (2007).
6. R. Yager. On ordered weighted averaging aggregation operators in multi criteria decision making. IEEE Transactions Systems Man and Cybernetics (1988). Vol. 18 pages 183-190.