Dictionary based hyperspectral retrieval

Miguel A. Veganzones[1], Mihai Datcu[2], Manuel Graña[1]

[1]Grupo de Inteligencia Computacional
Basque Country University (UPV/EHU), Spain

[2]German Aerospace Agency (DLR), Germany

ICPRAM 2012, Algarve (Portugal)

Outline

Motivation

- Compare hyperspectral images by Dictionary-based distances.
  - Approximation to the Normalized Compression Distance (NCD) based on compression factors -> computational expensive.
  - Grounded in the Normalized Information Distance (NID) based on Kolmogorov complexity -> uncomputable in a Turing sense.
- Apply to a Hyperspectral Content-Base Image Retrieval System (Hyperspectral CBIRS).
  - Experimental study in a real hyperspectral scene.

Kolmogorov complexity

- The *conditional Kolmogorov complexity* of a signal x given a signal y, $K(x|y)$, is the length of the shortest program running in an universal Turing machine, that outputs x when fed with input y.
- The *Kolmogorov complexity* of x, $K(x)$, is the length of the shortest program that outputs x when fed with the empty signal λ, that is, $K(x) = K(x|\lambda)$.

Normalized Information Distance

- The *information distance*, $E(x,y)$, is an universal metric distance defined as the length of the shortest binary program in a Turing sense that, from input x outputs y, and from input y outputs x.

$$E(x,y) = \max\{K(x|y), K(y|x)\} \tag{1}$$

- The *normalized information distance*, $NID(x,y)$, is defined as:

$$NID(x,y) = \frac{E(x,y)}{\max\{K(x), K(y)\}} \tag{2}$$

- Non-computable in a Turing sense due to $K(\cdot)$.

Normalized Compression DIstance

- The *normalized compression distance*, $NCD(x,y)$, is a computable version of (2) based on a given compressor, C.

$$NCD(x,y) = \frac{C(xy) - \min\{C(x), C(y)\}}{\max\{C(x), C(y)\}} \quad (3)$$

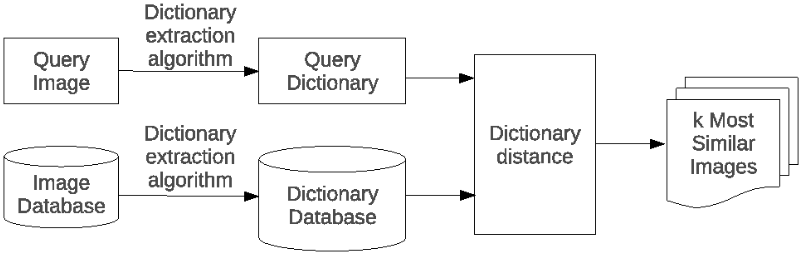
where $C(\cdot)$ is the length of a compressed signal by using compressor C, and xy is the signal resulting of the concatenation of signals x and y.

- If the compressor C is normal, then the NCD is a quasi-universal similarity metric.
- Computationally expensive due to $C(xy)$.

Dictionary distances

- Distances based on the codewords of the dictionaries extracted by means of dictionary-based compressors, such as the LZW for text strings.
- This dictionary approach only requires set operations to calculate the distance between two signals given that the dictionaries have been previously extracted.
- Thus, dictionary distances are suitable for mining large image databases where the dictionaries of the images in the database can be extracted off-line.

Proposed Hyperspectral CBIR system



Dictionary-based compression

- Given a signal x, a dictionary-based compression algorithm looks for patterns in the input sequence from signal x.
- These patterns, called *words*, are subsequences of the incoming sequence.
- The compression algorithm result is a set of unique words called *dictionary*.
- The dictionary extracted from a signal x is hereafter denoted as $D(x)$, with $D(\lambda) = \emptyset$ only if λ is the empty signal.
- The union and intersection of the dictionaries extracted from signals x and y are denoted as $D(x \cup y)$ and $D(x \cap y)$ respectively.

Dictionary distances

- Normalized Dictionary Distance (NDD):

$$NDD(x,y) = \frac{D(x \cup y) - \min\{D(x), D(y)\}}{\max\{D(x), D(y)\}}, \tag{4}$$

- Fast Dictionary Distance (FDD):

$$FDD(x,y) = \frac{D(x) - D(x \cap y)}{D(x)}. \tag{5}$$

- NDD and FDD are both normalized admissible distances satisfying the metric inequalities.
- Thus, they result in a non-negative number in the interval $[0,1]$, being zero when the compared files are equal and increasing up to one as the files are more dissimilar.

Outline

1. Introduction
2. Dictionary distances
3. Experiments
   - Experimental data
   - CBIR performance meassures
   - Experimental methodology
4. Results
5. Summary

HyMap scene

- The hyperspectral HyMAP data was made available from HyVista Corp. and German Aerospace Center's (DLR) optical Airborne Remote Sensing and Calibration Facility service.
- The sensed scene corresponds to the radiance captured by the sensor in a flight line over the facilities of the DLR center in Oberpfaffenhofen (Germany) and its surroundings, mostly fields, forests and small towns.
- The data cube has 2878 lines, 512 samples and 125 bands; and the pixel values are represented by 2-bytes signed integers.

HyMap patches and categories

- We cut the scene in patches of 64×64 pixels size for a total of 360 patches.
- We grouped the patches by visual inspection in five rough categories:
  - The three main categories: 'Forests', 'Fields' and 'Urban Areas'.
  - A 'Mixed' category for those patches that presented more than one of the three main categories.
  - A 'Others' category for those patches that didn't represent any of the above or that were not easily categorized by visual inspection.
- The number of patches per category are: (1) Forests: 39, (2) Fields: 160, (3) Urban Areas: 24, (4) Mixed: 102, and (5) Others: 35.

HyMap datasets

- We defined three datasets:
  - Dataset 1: patches belonging to the three main categories: Forests, Fields and Urban Areas.
  - Dataset 2: we added patches from the fourth category: Mixed.
  - Dataset 3: contains the patches from all five categories.
- Datasets present an increase in complexity.

Outline

Precision-Recall curves

- Precision, p, is the fraction of the returned images that are relevant to the query.
- Recall, q, is the fraction of returned relevant images respect to the total number of relevant images in the database according to *a priori* knowledge.
- If we denote T the set of returned images and R the set of all the images relevant to the query, then

$$p = \frac{|T \cap R|}{|T|} \; r = \frac{|T \cap R|}{|R|}$$

- Precision and recall follow inverse trends when considered as functions of the scope of the query.
  - Precision falls while recall increases as the scope increases.

Average PR curves

- To evaluate the overall performance of a CBIR system, the Average Precision and Average Recall are calculated over all the query images in the database.
- For a query of scope k, these are defined as:

$$\bar{p}_k = \frac{1}{N} \sum_{\alpha=1}^{N} p_k(H_\alpha) \tag{6}$$

$$\bar{r}_k = \frac{1}{N} \sum_{\alpha=1}^{N} r_k(H_\alpha). \tag{7}$$

Normalized Rank

- The Normalized Rank is a performance measure used to summarize system performance into an scalar value.
- The normalized rank for a given image ranking Ω_α, denoted as $\mathrm{Rank}(H_\alpha)$, is defined as:

$$\mathrm{Rank}(H_\alpha) = \frac{1}{NN_\alpha}\left(\sum_{i=1}^{N_\alpha} \Omega_\alpha^i - \frac{N_\alpha(N_\alpha - 1)}{2}\right), \quad (8)$$

  where N is the number of images in the dataset, N_α is the number of relevant images for the query H_α, and Ω_α^i is the rank at which the i-th image is retrieved.
- This measure is 0 for perfect performance, and approaches 1 as performance worsens, being 0.5 equivalent to a random retrieval.

Average Normalized Rank

- The average normalized rank, *ANR*, for the full dataset is given by:

$$ANR = \frac{1}{N} \sum_{\alpha=1}^{N} \text{Rank}\left(H_{\alpha}\right). \tag{9}$$

Outline

Experiments

- We independently test the NCD, the NDD and the FDD distances in three CBIR experiments corresponding to each of the three previously defined datasets.
- Each hyperspectral image is first converted to a text file in two ways: pixel-wise and band-wise.
- The NDD and FDD are calculated using the dictionaries extracted by the LZW compression algorithm.
- The NCD is calculated by CompLearn software using default options, that is BZLIB compressor.

CBIR experiment

- For each hyperspectral image H_α in a dataset we calculate the dissimilarity measure between H_α and each of the remaining images in the dataset using a selected distance.
- These dissimilarities are represented as a vector $\mathbf{s}_\alpha = [s_{\alpha 1}, \ldots, s_{\alpha N}]$, where N is the number of images in the dataset and $s_{\alpha,\beta}$ is the dissimilarity between the images H_α and H_β, with $\alpha, \beta = 1, \ldots, N$.

CBIR experiment (II)

- We can define the ranking of the dataset relative to the query image, $\Omega_\alpha = [\omega_{\alpha,t} \in \{1,\ldots,N\}; t = 1,\ldots,N]$, as the set of image indexes ordered according to increasing values of their corresponding entries in the dissimilarity vector $\mathbf{s}_\alpha$.
- That is, we sort in increasing order the components of $\mathbf{s}_\alpha$, and the corresponding rendering of image indexes constitute Ω_α, so that $s_{\alpha,\omega_{\alpha,t}} \leq s_{\alpha,\omega_{\alpha,t+1}}$.

CBIR performance

- For each hyperspectral image H_α, a query $Q_k(H_\alpha)$ is formulated returning the k most similar (less dissimilar) images H_β in the dataset relative to the image H_α, where k is the scope of the query and takes values in the range $1 \leq k \leq N$.
- The groundtruth for a query image H_α is a ranking, Ω_α^{GT}, given by the a-priori categorization made by visual inspection.
- This way, the relevant set for a query patch H_α is formed for all those patches belonging to its same category $\mathscr{C}(\alpha)$.

CBIR performance (II)

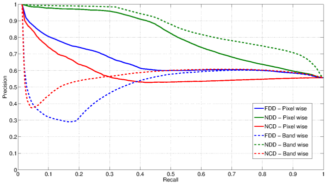
- Given a query $Q_k(H_\alpha)$, the set of returned images $T_k(H_\alpha)$ and the set of relevant images $V_k(H_\alpha)$ are defined as follows:

$$T_k(H_\alpha) = \Omega_{\alpha,k} = \left[\omega_{\alpha,p} \text{ s.t. } s_{\alpha,\omega_{\alpha,p}} \le s_{\alpha,\omega_{\alpha,k}}\right] \tag{10}$$

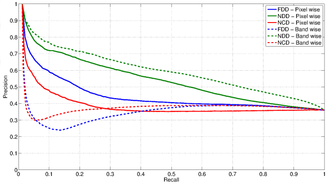
$$V_k(H_\alpha) = \Omega_\alpha^{GT} = [\beta \text{ s.t. } \mathscr{C}(\beta) = \mathscr{C}(\alpha)] \tag{11}$$

  where $\mathscr{C}(\gamma)$ indicates the category to which the patch H_γ belongs.
- Now $T_k(H_\alpha)$ and $V_k(H_\alpha)$ can be used to calculate the average precision and recall measures of the system, as well as the average normalized rank.

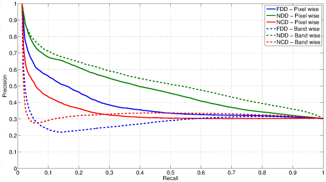
Precision-Recall curves



Experiment 1



Experiment 2



Experiment 3

PR curves conclusions

- In all the experiments NDD outperforms the other distances independently of the image to text string conversion ordering used.
- NCD outperforms FDD showing that the lack of a normalization factor in the FDD is an important issue.
- The band-wise NDD gives the best performance in all the experiments, due to the high correlation on consecutive bands.
- However, surprisingly, the band-wise ordering shows a bad performance for low recall values using FDD and NCD, improving as the recall values increase up to performances similar to the pixel-wise ordering.

PR curves conclusions (II)

- In general, the performance decreases smoothly as we include hardest categories, yielding still good precision-recall values for the NDD function.
- Also, NCD presents a general lower precision compare to dictionary-based distances, although its performance decreases more slowly than the performances of NDD and FDD as we add more difficult categories.

Average Normalized Rank

Experiment 3

	ANR		
Category	**FDD**	**NDD**	**NCD**
Forests	0.065	**0.049**	0.162
Fields	0.323	**0.235**	0.315
Urban Areas	**0.011**	0.013	0.107
Mixed	**0.246**	0.254	0.318
Others	**0.197**	0.232	0.425
Average	0.169	**0.156**	0.266

Pixel-wise ordering

	ANR		
Category	**FDD**	**NDD**	**NCD**
Forests	0.130	**0.061**	0.304
Fields	0.369	**0.164**	0.226
Urban Areas	0.006	**0.008**	0.674
Mixed	0.254	0.250	0.360
Others	**0.177**	0.210	0.570
Average	0.187	**0.139**	0.427

Band-wise ordering

ANR conclusions

- ANR results confirms the average outperform of NDD over FDD and NCD, although FDD slightly outperforms NDD in some cases.
- Interestingly, ANR can partially explain the effect in the FDD and NCD precision-recall curves using band-wise ordering for low recall values:
  - FDD is having problems retrieving the 'Fields' category.
  - NCD is having problems retrieving the 'Forests' and 'Urban Areas' categories.
- Further experiments must be conducted to give a better explanation to why band-wise ordering affects so much FDD and NCD performance.

Summary

- We have introduced a hyperspectral CBIR System using dictionaries.
- The dictionaries approach solves the computational cost problem by approximating NCD using dictionaries extracted offline from each of the database images.
- Results using real hyperspectral datasets show that the NDD outperforms the FDD and the NCD.
- In order to extract the dictionaries (or compress the signals for the NCD) the arrangement of the image data affects severelly the performance of the FDD and NCD similarity functions.
  - Further experiments must be conduced to find an explanation of that unexpected effect.
- Generally, we can conclude that the presented results validate the use of dictionaries for hyperspectral image retrieval.

Thanks for your attention

- Miguel A. Veganzones: miguelangel.veganzones@ehu.es
- Prof. Manuel Graña: manuel.grana@ehu.es
- Prof. Mihai Datcu: mihai.datcu@dlr.de
- Grupo de Inteligencia Computacional: www.ehu.es/computationalintelligence