Browsing through and searching for similar images in astronomical archives using data density-based image icons

André Csillaghy

Institute of Astronomy, ETH-Zentrum,
CH-8092 Zurich, Switzerland
csillag@astro.phys.ethz.ch

Abstract

This article addresses the issue of retrieving images from large astonomical archives. It presents a method to define indexing features describing specific characteristics of the information contained in the image. Indexing features allow to compute a "degree of similarity" between images. In the method presented here, indexing features are derived from image icons. The latter represent symbolically the image content and are mainly used for browsing. The transition from icons to indexing features is done using a self-organizing map (SOM). In image retrieval systems, SOM-generated indexing features allow to reach high levels of retrieval precision. This is illustrated with ASPECT, a system managing the Zurich archive of solar radio spectrograms. For specific queries and for recalls less than 10%, a precision above 50% have been reached. It represents about 20% increase compared with a retrieval system based on global indexing features.

1 Overview

The efficiency and effectiveness of a retrieval system for large image archives relies on two main actions:

- *Browsing through a large number of images* allows to visualize roughly but quickly the contents of the archived images. To be quick, browsing uses lossy compressed versions of the images or symbolic image descriptions like image icons (Csillaghy, 1994).
- *Searching for similar images* allows to select the set of images to browse through. This implies the determination of some "degree of similarity" between images. The degree of similarity relies on *indexing features* that describe specific characteristics of the structures contained in the images.

Traditional methods to define indexing features rely, for instance, on text association (Murtagh, 1994), color histograms (Flickner et al., 1995), low-level image properties (Gupta and Jain, 1997) or texture description (Carson et al., 1997). To process a large number of images, the actual values of the indexing features associated with each document must be derived automatically. The automatization is problematic. Traditional methods are usually developed either for conventional photographic pictures (press photographs, museum catalogues etc.) or for small collections of images. Astronomical images archives, on the other hand, are large. Moreover, the archived images are usually noisy and mostly contain diffuse structures. Generally, the methods mentioned above cannot be applied to them.

To define indexing features for astronomical images, another approach must be used. The method presented here uses the information contained in image icons. An icon is composed of a set of *boxes* describing regions of similar texture (Section 2). Boxes are analyzed with a self-organizing map (Kohonen, 1995), which classifies the latter by shapes and volumes. This classification tells about the type of structures contained in the image (Section 3). The utilization of SOMs and icons to browse through and search for similar images is applied to the management of an archive of solar radio spectrograms (Section 4) using a system called ASPECT (ASPECT, 1996). The effectiveness of

the ASPECT system based on SOM-generated indexing features, is significantly increased compared with its effectiveness based on global indexing features (Section 5).

2 Transformation of images into icons

The method to transform image into icons have been described in details in another paper (Csillaghy, 1996). Its basic elements are recalled here.

A way to segment the image into domains with similar texture is investigated. To this end, image pixels are represented as points in a 3-dimensional attribute space, where two dimensions are given by the image axes and the third dimension is given by the color values of the parameterized pixels. The attribute space is partitioned into 3-dimensional *regions* that can hold only up to a fixed number of points. Due to this constraint, the regions have different shapes and volumes to adapt to the variable density of points in the attribute space. Consider the points in a given region. They deliver information about the local data distribution, and can be used to define a structure called a *box*. The latter represents the part of the image with similar texture.

Each individual box can be considered as a 6-tuple. Consider the r-th box, b_r, of an icon. Three values determine its position in each dimension. They are determined by the average of the points in the associated region,

$$\mu_{r,i} = \frac{1}{m_r} \sum_{k=0}^{m_r-1} x_{k,r,i} \qquad (i = 1, 2, 3), \tag{1}$$

Where m_r is the number of points in the r-th region, $x_{k,r,i}$ is the value of the k-th point in the i-th dimension and in the region r. The remaining three values determine its extension in each dimension. They are determined by the standard deviations of the points in the associated region,

$$\sigma_{r,i} = \frac{1}{m_r} \sqrt{\sum_{k=0}^{m_r-1} (x_{k,r,i} - \mu_{r,i})^2} \qquad (i = 1, 2, 3). \tag{2}$$

Hence, the box representing the r-th region is written as

$$b_r = < \mu_{r,1}, \mu_{r,2}, \mu_{r,3}, \sigma_{r,1}, \sigma_{r,2}, \sigma_{r,3} > . \tag{3}$$

The set of boxes created following the method summarized above represents an abstraction of the full image information content. For a specific application, however, only a fraction of this information is interesting. Therefore, a selection procedure is used to determine which boxes actually represent the information wanted. The selection of boxes constitutes the image icon; it is visualized by projecting the selected boxes into the plane of the image (see Figure 1).

Image icons are usually displayed in a browser, for example Netscape, as illustrated in Figure 2. They occupy only a small area on screen, thus allowing to display many of them simultaneously.

3 Derivation of indexing features from icons

3.1 General considerations

Two images are compared by determining their "degree of similarity." This implies the definition of meaningful indexing features which codify the main characteristics of the structures in the image. Indexing features can be arranged in a specific order, thus building for each image a *document description vector*. The images most similar to a given reference image are those that have their associated document description vector nearest to the document description vector of the reference image. The notion of "nearest" will be further discussed in Section 3.2.

How can indexing features for astronomical images be adequately defined? The information contained in image icons can help. Using icons as source of indexing features is attractive because the information they contain has already been selected and abstracted:

- *Selected:* The boxes building the image icons have been selected as representing interesting information; thus, the computation of the values of indexing features is not biased by irrelevant data. For example, background or disturbances, which are a significant part of the original images, do not influence the computation.

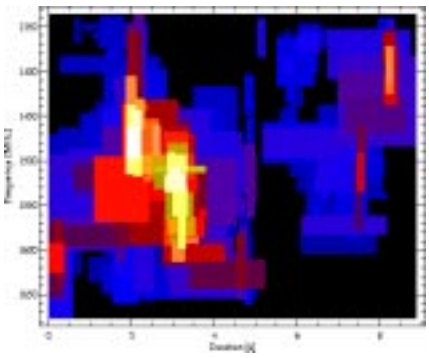

Figure 1: *Comparison between an image and its associated image icon. The image shows a solar radio spectrogram. It represents the density of the solar radio flux in the frequency-time plane. Enhanced emission is shown bright.*

– *Abstracted:* The information in icons is represented by simple 6-tuples that are easy to handle. The definition of indexing features is therefore relatively straightforward. Moreover, using icons for the computation of indexing features does not require any additional access to the large-size original documents.

Many ways can be used to define indexing features from icons. A number of *global indexing features* have been defined in another paper (Csillaghy, 1997). Furthermore, *local indexing features* can be defined using, for instance, minimal spanning trees (Murtagh and Heck, 1987), principal component analysis (Jolliffe, 1986) or self-organizing maps. The rest of the paper focuses on the utilization of the latter.

3.2 Self-organizing maps for the definition of indexing features

3.2.1 Principles of self-organizing maps

Basically, a self-organizing map (SOM) is a two-dimensional artificial neural network, that is, an array of interconnected cells. It has been described by Kohonen (1995) and is schematized in Figure 3. The spatial location of a cell in the map corresponds to a specific region of the multidimensional attribute space to be analyzed, or *input space*. A *cell* of the map reacts—that is, switches from its unactivated state 0 to its activated state 1—when a data point of the input space "presented" to the map originates from the corresponding region of the input space. The correspondence between regions of the input space and cells of the map is determined during a training process.

A set of data samples are used to train the map. During training, the map has its cells i, characterized by reference vectors m_i that are updated after each input. The learning process is described by:

$$m_i(t+1) = \begin{cases} m_i(t) + \alpha(t)[x(t) - m_i(t)] \\ \qquad\qquad \text{if} \quad i \in N(t) \\ m_i(t) \qquad\qquad \text{otherwise,} \end{cases} \tag{4}$$

where t is the (discrete) updating time, $x(t)$ is the current input vector and $\alpha(t)$ is called the *adaptation gain*, $0 < \alpha(t) < 1$. The *neighborhood function* $N(t)$ determines which cells must be updated for a given input point. Thus, the topological structure of the input space is conserved in the map.

Once the training phase is completed, the reference vectors are left unchanged. They determine which cell must react to an (now arbitrary) input point. Since the topology of the input space is conserved by the map, nearby cells react also to nearby input classes. The conservation of the topology allows to visualize a multidimensional space in two dimensions. Therefore, maps can be used to visualize the differences between input points occuring in a data set.

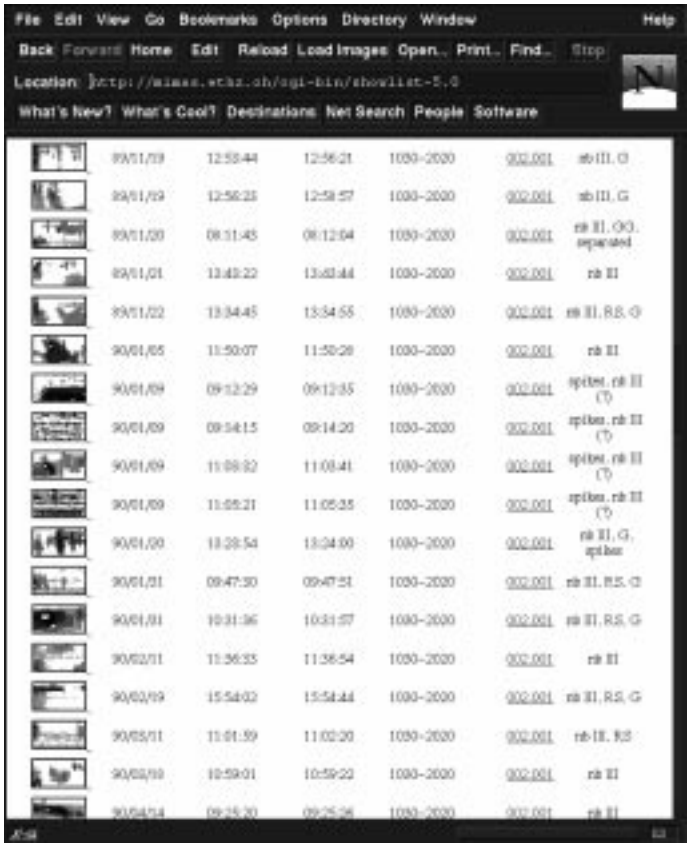

Figure 2: *This browser displays a list of images, ususally in response to a query. Here,* narrowband type III bursts (nb III) *have been requested. The information displayed consists of the following elements (from left): the (clickable) image icon, the date, the start and end time of the observation, the frequency range (in MHz), the frequency program number and remarks about the event (including the burst types).*

3.2.2 Self-organizing maps and boxes

Figure 4 illustrates how a SOM works with boxes. It is first trained to react to the shape of boxes. Then, as shown in the figure, two samples are "presented" to the trained SOM. A box of a given shape generates a cell reaction at the bottom left of the map. The other box, which has a significantly different shape, generates a cell reaction at the top right. Remarkably, if the shape of two boxes differ only slightly, their corresponding cell reaction are also nearby in the map.

The boxes contained in image icons are used as input data to SOMs. Each box produces a single cell reaction. By summing the cell reactions corresponding to individual boxes, a "total" map can be associated with a given icon. This approach is attractive for the following reasons:

- About 300 boxes are contained in a single icon. Thus, a large quantity of boxes is available in the whole archive. This allows to train the SOM with a large number of samples.
- Because a map represents a sum of cell reactions, and not only a single cell reaction, local properties of the image content can be revealed.

Indexing features are defined by associating a cell of the SOM with each single indexing feature.

3.3 Measure of the similarity between documents

Remember that indexing features are arranged in a specific order, thus building a document description vector. The distance between description vectors, in a given metric, corresponds to the similarity

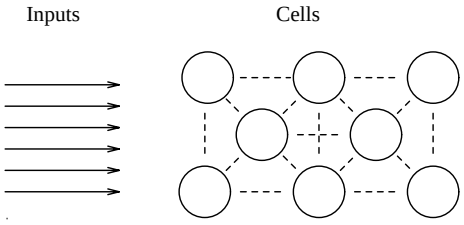

Figure 3: *A self-organizing map consists of an array of interconnected cells. A region of the input data space is associated with each cell of the network. The dimensionality of the input space is arbitrary.*

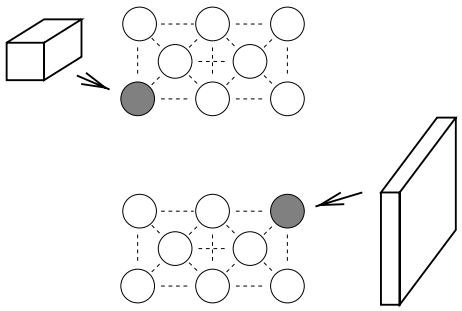

Figure 4: *A simple functional example of the self-organizing map. Assume that the map is trained to react to boxes (characterized by 6-tuples). It will organize itself so that different shapes of boxes will make different regions of the map react.*

between documents. Consider the space spanned by the description vectors, called the *description vector space*. Consider also two document description vectors, $\vec{d_j}$ and $\vec{d_k}$ in the description vector space. Their distance can be measured, for instance, using the Euclidean norm:

$$\rho_E(\vec{d_j}, \vec{d_k}) = ||\vec{d_j} - \vec{d_k}|| = \left(\sum_{i=0}^{m_\varphi - 1} (\vec{d_j} - \vec{d_k})^2 \right)^{\frac{1}{2}}. \tag{5}$$

However, ρ_E is not necessary the best measure of the "degree of similarity." It fails, for instance, if similar documents are aligned in a specific direction of the description vector space. In this case, the distance function should take into account the non-isotropic distribution of points in the description vector space.

For this purpose, another function is often used (mainly in the context of textual information retrieval), is the direction cosine:

$$\cos(\theta_{j,k}) = \frac{\vec{d_j} \cdot \vec{d_k}}{||\vec{d_j}|| \quad ||\vec{d_k}||}. \tag{6}$$

$\cos(\theta_{j,k})$ assumes a linear dependence between similar document description vectors. The case $\cos(\theta_{j,k}) = 1$ represents the highest correlation between $\vec{d_j}$ and $\vec{d_k}$, and therefore corresponds to the highest similarity.

4 Applications to solar radio spectrograms

Solar radio spectrograms consist of images, often called *events*, which display the density of the solar radio flux in the time-frequency plane. Spectrograms share the typical characteristics of astronomical images: they have a low signal-to-noise ratio, and they contain diffuse structures. The structures are divided into types and sub-types. They correspond to signatures of the emision produced by the

Main type	$\lvert\mathcal{D}^{\text{rel}}_{\text{main}}\rvert$	Subtype	$\lvert\mathcal{D}^{\text{rel}}_{\text{sub}}\rvert$
III	265	narrowband	114
		broadband	73
		large group (> 5 bursts)	26
		small group (< 5 bursts)	21
		reverse drift	31
IV	56	modulated	15
		fibers	33
		zebras	8
Blips	32	III-like	11
		patchy	21
Pulsations	58		58
Patches	39	cloudy	14
		cigar	7
		caterpillars	5
		large spots	13
Spikes	50		50

Table 1: *The classification of radio bursts in types and subtypes used for the evaluation of the ASPECT system A test collection $\mathcal{D}$ of 437 images is used. The number of images per main type and sub-type classes are given by $\lvert\mathcal{D}^{\text{rel}}_{\text{main}}\rvert$ and by $\lvert\mathcal{D}^{\text{rel}}_{\text{sub}}\rvert$, respectively.*

acceleration of high-energy particle in the solar corona. The list of types and subtypes used in this work is given in Table 1.

Parameter	Ordering phase	Fine tuning phase
Number of training samples	10,000	100,000
Radius of the neighborhood function	30	3
Adaptation factor	0.9	0.02

Table 2: *The parameters used to train the SOM.*

4.1 Parameters for training the SOM

The SOM is trained in two phases. First, the *ordering phase* determines the values of the reference vectors, to establish roughly the correspondence between the topology of the input space and the map. Second, the *fine tuning phase* adjusts accurately the values of the already ordered reference vectors.

The SOM package used in this work has been developed by the group of Kohonen (SOM, 1995). For the tests presented below, a map of a dimension of 30x30 was used. The parameters given below are determined experimentally. The are either determined specifically for ASPECT, or for SOMs in general. They are summarized in Table 2.

4.2 Maps for a single type of radio emission

Figure 5 compares maps associated with three images of the same burst type. These images belong to the class of "type III bursts." The maps present the following characteristics:

- There are two main regions where reactions are registered. The first region is around $(15, 15)$. The second region is at $(20 - 25, 5 - 10)$ and is more spreaded.
- Some regions of the map recorded no reactions at all.

– Some regions of the map recorded a relatively small number of reactions, but for each map at different locations

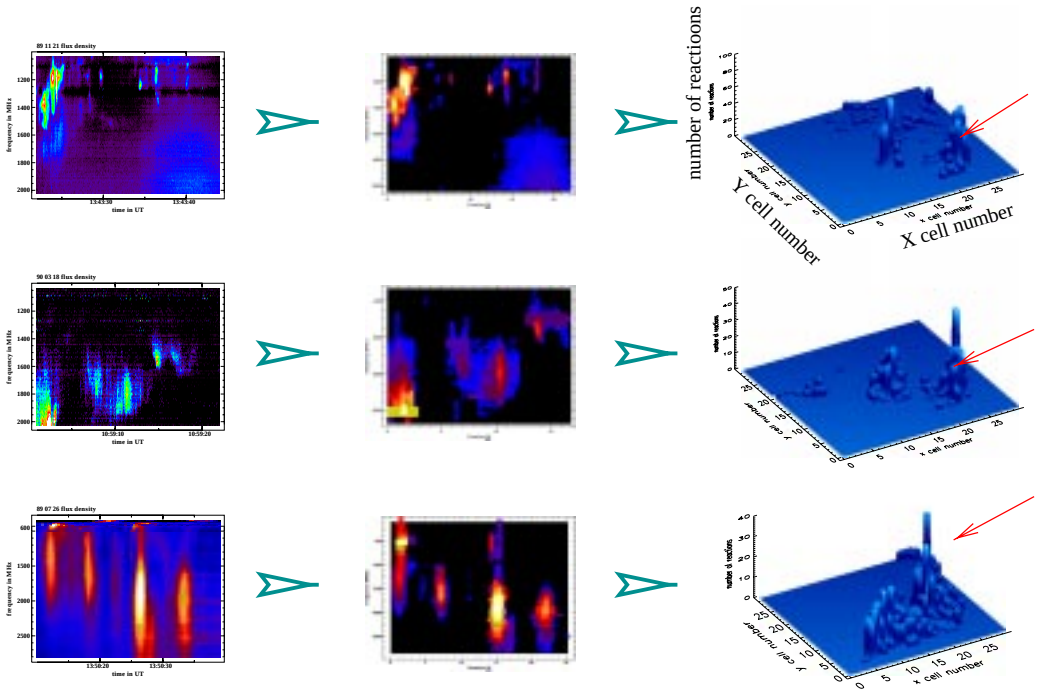

Figure 5: *Three spectrograms of the same type of radio burst. The map reactions are located in the same regions (arrows). Other regions did not react at all.*

4.3 Map for a different types of radio emisson

Figure 6 compares maps associated with three images of different burst types. These classes are: (1) type III bursts, (2) millisecond radio spikes and (3) type IV bursts. The maps present the following charateristics:

– There is no correlation between the type III-burst map and the type IV-burst map. This corresponds to expectations: type-III and type-IV bursts correspond to signatures of different processes.

– Between the type III-burst map and the millisecond-spikes map, only a few cells have reactions in common.

– Between the type IV burst-map and the millisecond-spikes map, there is a correlation in the region $(10-20, 25-30)$. For type IV however, a whole region of the map have reacted with no correlation with the other maps.

5 Retrieval effectiveness

An image retrieval system can be evaluated by considering its capacity to effectively retrieve information relevant to a user. It is called the *retrieval effectiveness*. Below, it is measured for ASPECT. The indexing features considered are derived using the method presented in the previous section.

The retrieval effectiveness is measured by the *recall* and the *precision* (van Rijsbergen, 1979). For a given query and a given number of documents retrieved, the recall gives the ratio between the

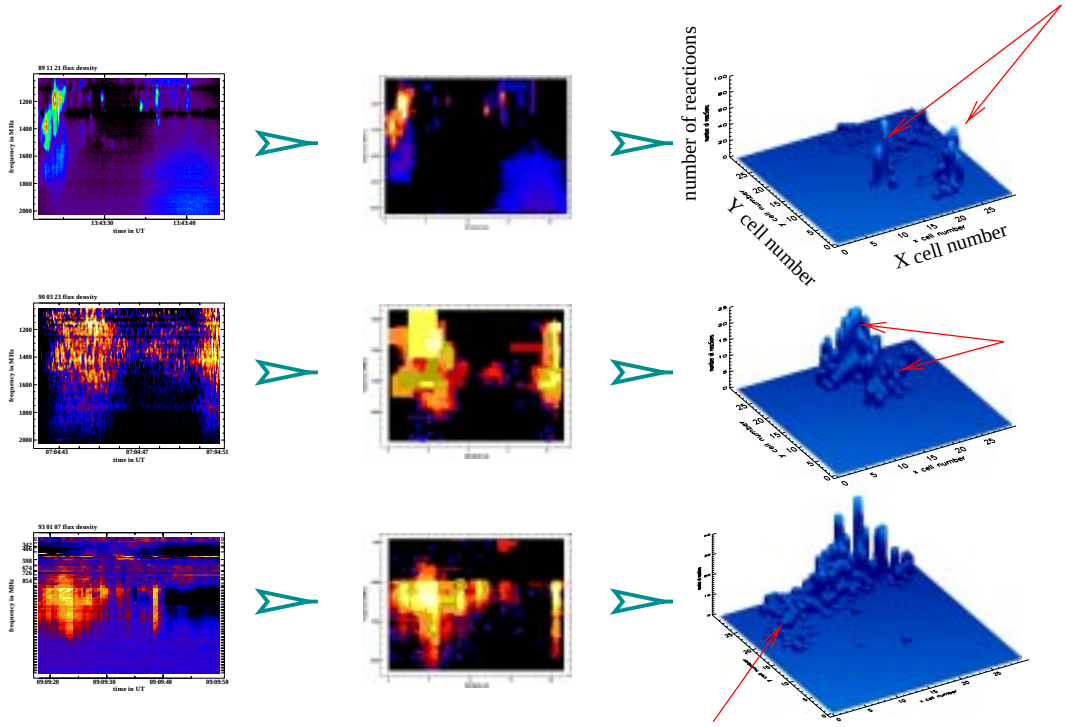

Figure 6: *Three spectrograms of different types of radio bursts. The map reactions are located in different domains (arrows). For instance there is little correlation between the first map (type III burst) and the third map (type IV burst).*

number of relevant documents retrieved and the total number of relevant documents in the collection considered. The precision gives the ratio between the number of relevant documents retrieved and the number of retrieved documents.

Recall and precision values for a system can be represented in a *recall and precision graph* (Frei et al., 1991; Raghavan et al., 1989), where the precision of the system is plotted as a function of the recall. This representation allows, for instance, to compare the effectiveness of different retrieval functions. The method to derive a recall and precision graph on the basis of these two values is described by Schäuble (1997).

The recall and precision graph for ASPECT is computed as follows. 32 reference ("query") images are selected from a test collection of 437 images. They contain solar radio bursts that are divided into several *main types* and *subtypes* (see Table 1). The classification processed by the retrieval system is compared with a classification that have been done by hand (Isliker and Benz, 1994). In this way, it is possible to decide if an image is relevant or not.

Two reference images are selected per subtype. For these images, a search for similarity is started. The SOM-based response of the system is compared with a global indexing feature-based response (Csillaghy, 1997). The similarity between the reference and the other images is computed using the Cosine function given in Equation 6. The resulting recall and precision graph is shown in Figure 7.

The graph shows that the precision is better for SOM indexing features. For low recalls, the precision is high: when considering the classes of type III bursts and type IV bursts, a precision above 50% for recalls lower than 10% is attained. Unfortunately, the precision breaks down if classes with less elements are considered. Moreover, sub-classses retrieval precision is also much lower. Nevertheless, the SOM-generated indexing features lead to a better precision than global indexing features.

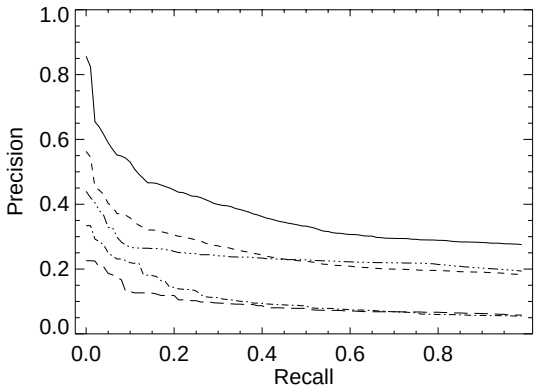

Figure 7: *The recall and precision graph for the ASPECT retrieval system. Retrieval status values are computed with the Cosine function. The following graphs are compared: Solid line: retrieval of type III and type IV bursts only. Dashed line: retrieval of all main types. Dashed-dotted line: retrieval of all subtypes. Dashed-triple-dotted line: retrieval of main types with global indexing features. Long-dashed line: retrieval of subtypes with global indexing features.*

6 Conclusions

We have shown that SOMs can be used to attain high levels of retrieval effectiveness. Methods to improve the precision, especially when considering subtypes have to be further investigated. This can be done for instance by giving more or less weight to some boxes (or to some regions of the map) when analyzing the image icons.

Acknowledgements

The author acknowledges helpful discussions with A.O.Benz, H.Hinterberger and P.Schäuble.

References

ASPECT, 1996, *The Ikarus/Phoenix Image Retrieval System ASPECT*, ETH Zurich, http://mimas.ethz.ch/archive.html

Carson, C., Belongie, S., Greenspan, H., and Malik, J., 1997, in *Proc. Workshop on content-based access of image and video libraries*, Vol. 4, IEEE

Csillaghy, A., 1994, in *Proceedings of the First Int. Conf. on Image Processing*, IEEE Computer Society Press, Los Alamitos, http://www.astro.phys.ethz.ch/papers/csillaghy/

Csillaghy, A., 1996, *Vistas in Astronomy* **40**, 503, http://www.astro.phys.ethz.ch/papers/csillaghy/nice.ps

Csillaghy, A., 1997, *Proc. of the 5th Int. Workshop on Data Analysis in Astronomy*, in press, http://www.astro.phys.ethz.ch/papers/csillaghy/

Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Patrovic, D., Steele, D., and Yanker, P., September 1995, *Computer* pp 23–32

Frei, H., Meienberg, S., and Schäuble, P., 1991, in N. Fuhr (ed.), *Workshop on information retrieval*, Vol. 289 of *Informatik-Fachberichte*, pp 1–10, Springer, Berlin

Gupta, A. and Jain, R., 1997, *Comm. ACM* 40(5)

Isliker, H. and Benz, A. O., 1994, *A&AS* **104**, 145

Jolliffe, I., 1986, *Principal component analysis*, Springer

Kohonen, T., 1995, *Self-Organizing Maps*, Springer, Berlin

Murtagh, F., 1994, in *Proceedings of the 14th Int. CODATA Conference*, Chambéry, http://http.hq.eso.org/~fmurtagh/papers/image-retrieval-codata94.ps

Murtagh, F. and Heck, A., 1987, *Multivariate Data Analysis*, Astrophysics and Space Science Library, D. Reidel, Dordrecht

Raghavan, V., Jung, G., and Bollman, P., 1989, *ACM Transactions on Information Systems* **7(3)**, 205

Schäuble, P., 1997, *Multimedia Information Retrieval: Content-Based Information Retrieval from Large Text and Audio Databases*, Kluwer, Dordrecht

SOM, 1995, *SOM_PAK: The Self-Organizing Map Program Package*, SOM Programming team, Helsinki University of Technology, http://nucleus.hut.fi/nnrc/som_pak/

van Rijsbergen, C., 1979, *Information retrieval*, Butterworth, London