Generic Image Structures in Integrated Media

B.N. Rossiter
School of Computing and Mathematics
University of Northumbria at Newcastle NE1 8ST, UK
nick.rossiter@unn.ac.uk; Tele: ++44 191 227 4662

M.A. Heather
University of Northumbria at Newcastle NE1 8ST
m.heather@unn.ac.uk; Tele: ++44 191 227 3034

Abstract

The generation, capture, storage, processing and retrieval of image-data objects requires a universal representation for the open-ended global environment of integrated media. That is a representation integrable at the content level both semantically and pragmatically with text based and multi-modal data types. Interoperability requires the same platform-independent methods to be applicable whatever the application. Open systems impose common *de facto* standards which possess an underlying naturality expressable in formal categories. It is argued that all image forms and their complex structures are representable in the same general pullback category.

Keywords: images, databases, generic structures, category theory.

1 Current State of Image Representation

There have been two main directions to date in handling image data: 1) Image contents modelled as a set of attributes at a fairly high level of abstraction but little scope for free or ad hoc queries [Jain and Dorai, 1997, Ogle and Stonebraker, 1995]. 2) Feature extraction/object recognition subsystems with automated object recognition but with methods that are difficult, computationally expensive and domain specific [Srihari 1995] as for instance in trademark recognition systems [Eakins 2000]. Rather than employ general models as used in databases (relational, network and hierarchical), information retrieval has traditionally employed customized methods (for example relevance matching or ranking algorithms). Text retrieval tends to be performed by specialised packages, usually involving inverted files. SPIRES was an early example of the use of a generalised database management approach to information retrieval which used quite an advanced form of dynamic data modelling which was particularly fitted to complex documents in full text [Heather and Rossiter, 1987]. Comparable database systems are now needed to store image data that can be searched with queries at the intensional level. Based on the experience of text retrieval, customized systems for similarity based image retrieval using only relevance feedback techniques of IR may be too simplistic unless they can extend to the extra layers of metadata to be found in database systems [Rossiter and Heather, 2000]. It is to be noted that there is a prominent use of text to assist searches in the systems based on attributes like Chabot [Ogle and Stonebraker, 1995]. Text is extensively used to support image retrieval because of the natural adjointness between text and image data.

The theoretical problems are: 1) to emphasise powerobjects rather than atomic objects with flexible searching on clusters and groups; 2) to construct universal relations for new connections intra-schema (local universe) and inter-schema (global universe) - i.e. for integration of images with text across heterogeneous databases; 3) to join type/domain attributes for different image representations (pixel, graph, Postscript). A number of set theoretic approcahes have been developed such as object-oriented methods but there are problems in unifying these [Barry and Stanienda, 1998].

Category theory is a very convenient mathematical workspace to bring together all the relevant techniques and the wide range of specialised methods and tools available. Category theory is able to unify many standard mathematical ideas which are needed in information processing [Barr and Wells, 1990] for a knowledge engineering context [Nelson and Rossiter, 1996] and in particular for objects like graphs, semantic nets, geometric models and hierarchies, as used in image work [Gebhardt 1997]. Categories provide a theory of types. Typing is an inherent feature of every image recognition with two basic categories of data, the source and the medium. Thus for an old master the source will be a human painter, whereas the medium may well be a painting in oils which will again import certain characteristics to the image and be specifiable in the retrieval process. A computer-generated image from automatic methods could have a source like natural physical processes in meteorology or natural biological processes in medical applications. These two examples might today very well be in some medium with similar characteristics. Both have bit streams but still have different categorial types for their respective sources. These full typing features need to be available to the system as necessary for storage, retrieval, display, compatibility and integration. There is an adjointness between the two categories as discussed below.

2 Current Application Areas

There is no limit to the variety of potential applications of data images. Typical examples of the importance of their detailed use can be seen in the speed sensor microwave imager for predicting long-term rainfall or in the special sensor microwave imager [Wetzel *et al* 1996]. In utilising the information to be derived from the images to predict long-term rainfall, there is a need to distinguish non-raining background conditions from any direct emission or scattering from the hydrometer. Similarly remote viewing with multi-spectral images allows 'nowcasting' of fog and pollution hazards but there is a need for image retrieval to take account of noise in either the source or the medium because thermal emission in the near infrared has to be eliminated. However, colour is often an important distinguishing feature. In the latter it can help to detect whether rain is present or not. A comparable need to extract information from the details of images and the corresponding rigorous requirements for precise query specifications are to be found in medical image applications. The information may be of an imprecise nature but knowledge extractable from the imprecision may be quite critical although not so easy to identify as in textual information.

A facet of information retrieval that does not feature much if at all in document retrieval is texture. These are perceptions of *qualia* [Ramachandran and Hirstein, 1997]. Notice the difference between the search for a portrait with a mature use of colour and a colour portrait of a mature face. Both involve qualia, the former at a syntactical level and the latter at a semantic level. Nevertheless despite difficulties, texture because of its importance is already being used as a search criterion. For instance with human faces, facial images have been shown to be distinguishable by a small number of relevant stimulus parameters based on the results of psycho-physical research. The perception of human texture is resolved into the three components of periodicity, directionality and randomness. [Picard and Minka, 1995] compare human texture with searching on texture of water in digital libraries. They suggest that no single model is sufficient. Some work on semantic interrogation has been done by [Yang and Wu, 1997] who propose a query-by-example diagrammatic language using type constructors like functions and inheritance to manipulate images at the semantic level.

3 Image Structures

One important concept derived from sheaf theory is the pullback where a product is restricted over some object or category. **S** and **M** both have arrows to some common category **W** (the real-world) as $\mathbf{S} \xrightarrow{\iota_s} \mathbf{W}$ and $\mathbf{M} \xrightarrow{\iota_m} \mathbf{W}$, and the subproduct of **S** and **M** over **IMG** written as $\mathbf{S} \times_{\mathbf{IMG}} \mathbf{M}$ may be represented by the diagram shown in Figure 1.

The diagram in Figure 1 describes the pullback of ι_m along ι_s. The product $\mathbf{S} \times_{\mathbf{IMG}} \mathbf{M}$ is an example of the universal limit. Holographic methods are examples of exploiting the concept of limit. It seems in general that the discovery of knowledge, as in information retrieval, is always the pullback of an arrow along another arrow over some category. The pullback limit is technically *left exactness*. This is the formal description of the existence of any knowable entity in the real world.

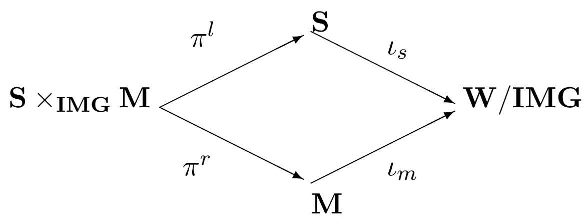

Figure 1: Pullback of types source ι_m along image ι_s

This example shows well the difference between the use of universal theory in constructive mathematics and the axiomatic set theory style of SQL where a kind of brute-force has to be applied to extract exact knowledge as a member of the powerset. The better scientific approach is to conceptualise from the three-level standpoint of this example. Thus all images are the limits in the pullback of two categories, a source (**S**) and a medium (**M**). The category **W** contains a subcategory **IMG** consisting of the real-world components that make up the source and the medium. These are real-world constructions. The important relationships between these indexed partial-orders are those of adjointness [Lawvere 1969] which lead to a formal understanding of the query. There is the need to find universal forms because they are natural; they should also be obvious and are usually quite trivial to identify.

The pullback diagram is actually richer than that shown in Figure 1. As indicated in Figure 2 many other arrows are in fact involved. Because of the principle of adjointness, these are unique which is why a particular arrow can imply that an image object exists and, if it exists, it can be retrieved by the appropriate query. The nature of these further arrows, together with those already introduced, is shown in the table in Figure 3.

4 Application to Image Processing

Consequently the pullback diagram in Figure 2 can be applied to the universal problem of image representation. The universal categories for any image are source **S** and medium **M**. For example a painter would be an object in the category **S** and oils an object in the category

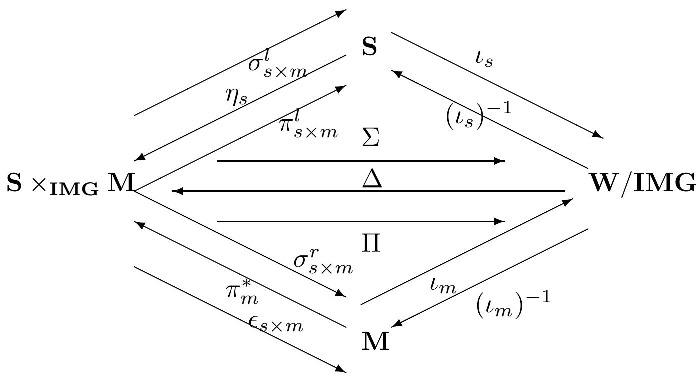

Figure 2: Pullback showing fuller collection of arrows

arrow	selects	of	from	comment
ι_s	W	given	S	source analysis
$(\iota_s)^{-1}$	S	given	W	source construction
ι_m	W	given	M	medium analysis
$(\iota_m)^{-1}$	M	given	W	medium construction
$\pi^l_{s\times m}$	S	given	$S \times M$	source nature
$\epsilon_{s\times m}$	M	given	$S \times M$	image *qualia*
Δ	$S \times M$	given	W	real-world image query
η_s	$S \times M$	given	S	image creativity
π^*_m	$S \times M$	given	M	medium type
$\sigma^l_{s\times m}$	S	some	$S \times M$	source collection
$\sigma^r_{s\times m}$	M	some	$S \times M$	medium collection
Σ	W	some	$S \times M$	component collection e.g. pixels
Π	W	all	$S \times M$	component combinations

Figure 3: Nature of each pullback arrow of Figure 2

M. For subcategory **S** painter and subcategory **M** for oils, the limit $\mathbf{S} \times_{\mathbf{IMG}} \mathbf{M}$ represents all the paintings by different painters and **W** includes particular components of the oil medium. For computer-generated images $\mathbf{S} \times_{\mathbf{IMG}} \mathbf{M}$ where the category **S** represents computers, **M** is the electronic medium and **W** consists of hardware items ranging from the components of the distributed IT systems to the pixel representations. That is the arrows, of the categories $\mathbf{S} \times_{\mathbf{IMG}} \mathbf{M}$ and **W**, each have a partial order self-indexed by their own semantics.

The arrow ι_s maps each source on to forms of representation. ι_m gives the forms used in a particular medium; this arrow performs the role of insertion of the category of medium into the physical or hardware form. The pullback (limit) $\mathbf{S} \times_{\mathbf{IMG}} \mathbf{M}$ contains all the components that make up the image forms. This is the generalised version of the familiar vector method and inner product. The limit $\mathbf{S} \times_{\mathbf{IMG}} \mathbf{M}$ is the outer product and the generalisation of the vector method and tensor products. In general the arrow
$\Delta : \mathbf{W} \longrightarrow \mathbf{S} \times_{\mathbf{IMG}} \mathbf{M}$
is the functor formally representing the discovery of knowledge in an operational sense. The table sets out the interpretation of the various arrows based on their universal form in category theory. Note that because of the three-level architecture, the highest type of arrow, natural transformation, can represent characteristics like creativity (η_s) or image quality ($\epsilon_{s\times m}$).

Systems that are open like modern distributed information systems are exposed to real-world complexity. Thus computer vision algorithms [Jain and Dorai, 1997] have from the outset had to accept the challenge of heterogeneous real-world systems. Category theory facilitates the integration of differeent models [Rossiter and Heather, 2000], thus assisting workers who use a selection of models to represent texture such as meta information, the hierarchical model and the object-oriented approach.

References

[Barr and Wells, 1990] Barr, M, and Wells, C, *Category Theory for Computing Science*, Prentice-Hall (1990).

[Barry and Stanienda, 1998] Barry, D, and Stanienda, T, Solving the Java Object Storage Problem, *IEEE Computer* **31**(11) 33-40 (1998).

[Eakins 2000] Eakins, J P, Trademark image retrieval - a survey, in: *Multimedia Storage and Retrieval Techniques - State of the Art* (Lew, M, ed). Springer-Verlag, Berlin (2000).

[Gebhardt 1997] Gebhardt, F, Survey on Structure-Based Case Retrieval, *Knowledge Engineering Review*, **12**(1) 41-58 (1997).

[Heather and Rossiter, 1987] Heather, M A, and Rossiter, B N, *Database techniques for text modelling: the document architecture of British statutes,* Computing Laboratory Tech Rep 227, 66pp, University of Newcastle upon Tyne (1987).

[Jain and Dorai, 1997] Jain, A K, and Dorai, C, Practising Vision: Integration, Evaluation And Applications, *Pattern Recognition*, **30**(2) 183-196 (1997).

[Lawvere 1969] Lawvere, F W, Adjointness in Foundations, *Dialectica* **23** 281-296 (1969).

[Nelson and Rossiter, 1996] Nelson, D A, and Rossiter, B N, Prototyping a Categorical Database in P/FDM. *Proceedings Second Int Workshop Advances Databases and Information Systems (ADBIS'95)*, Springer-Verlag Workshops in Computing, 432-456 (1996).

[Ogle and Stonebraker, 1995] Ogle, V E, and Stonebraker, M, Chabot: Retrieval from a Relational Database of Images, *IEEE Computer*, Sept. 40-48 (1995).

[Picard and Minka, 1995] Picard, R W, and Minka, T P, Vision Texture For Annotation, *Multimedia Systems*, **3**(1) 3-14 (1995).

[Ramachandran and Hirstein, 1997] Ramachandran, V S, and Hirstein, W, Three Laws of Qualia, *J Consciousness Stud* **4**(5/6) 429-457 (1997).

[Rossiter and Heather, 2000] Rossiter, B N, and Heather, M A, Handling Inconsistency with the Universal Reference Model, MS'2000 International Conference on Modelling and Simulation, Las Palmas de Gran Canaria, September 2000, 611-618 (2000).

[Srihari 1995] Srihari, R K, Use of Collateral Text in Understanding Photos, *Artif Intell Rev*, **8**(5-6) 409-430 (1995).

[Wetzel *et al* 1996] Wetzel, M A, Borys, R D, and Xu, L E, Satellite Microphysical Retrievals for Land-based Fog with Validation by Balloon Profiling, *J Appl Meteorology*, **35**(6) 810-829 (1996).

[Yang and Wu, 1997] Yang, L, and Wu, J K, Towards a Semantic Image Database System, *Data and Knowledge Eng*, **22**(2) 207-227 (1997).