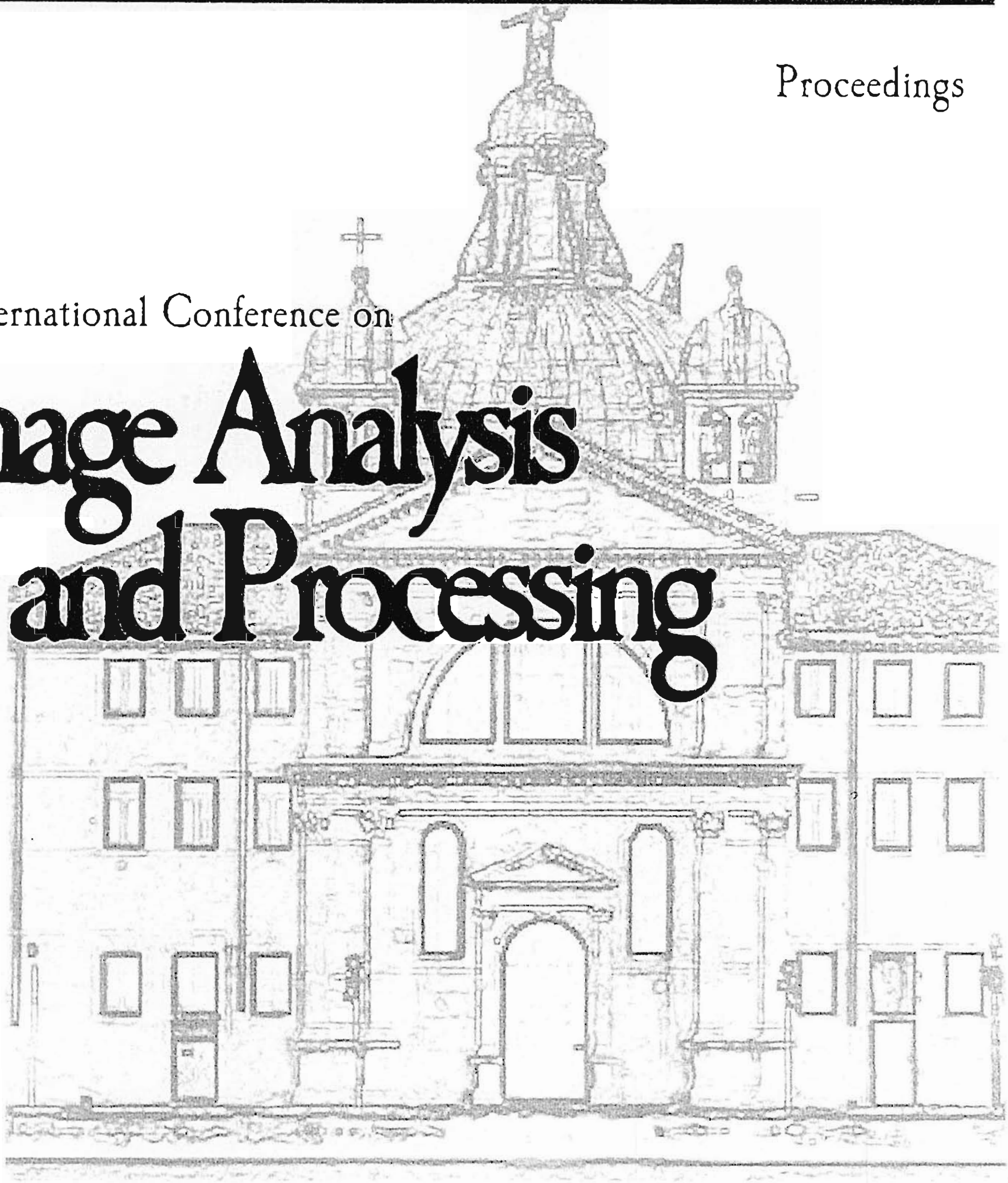
Proceedings

10th International Conference on Image Analysis and Processing



September 27–29, 1999
Venice, Italy

Sponsored by
International Association for Pattern Recognition
The University of Udine

Design of a Visual Environment for Evaluating and Customizing Medical Image Compression Techniques.

A. Signoroni, A. Bianchi, P. Mussio, R. Leonardi
Dipartimento di Elettronica per l'Automazione, Università di Brescia
{signoron, bianchi, mussio, leon}@bsing.ing.unibs.it

Abstract

This paper discusses the design of the Image Compression Laboratory (ICL), a visual environment supporting radiologists in interactively compressing medical images still maintaining the diagnostic information. In ICL lossy image compression techniques and interactive image interpretation approach are merged following the Cooperative Visual Environment approach. In this way the radiologists themselves can perform the compression only relying on their professional skill and knowledge.

1. Introduction

End-users, experts in some field but not in computer science, image processing and pattern recognition [1], increasingly use these techniques to perform critical tasks, of which they are legally responsible [2]. For example radiologists have the responsibility of managing the quality of images and recognizing in them medical structures to derive their diagnosis and prognosis [3]. They also need to archive these images and diagnosis as well as to distribute them to other colleagues. When using computer based systems, radiologists need to maintain the quality of images at the level required for their tasks, but also need to compress them (possibly degrading the image quality) due to a) the huge amount of data generated by digital imaging systems and b) the delay time which is often prohibitive for many interactive activities (e.g. tele-diagnostics). To face these contrasting requests, radiologists rely on tools based on advanced mathematical techniques, that often are not familiar to them. In friendly systems, it becomes easy the use of the tools but remains difficult for a radiologist to steer the tuning of the tools to specific case and to evaluate the practical implications of the parameter settings on quality (precision, accuracy and reliability) of the results.

This paper discusses the design of the Image Compression Laboratory (ICL), a visual environment in which end-users process and manage images to be distributed to their colleagues. ICL is developed according to the *Cooperative Visual Environment* (*CVE*) [4, 5] approach and architecture. A CVE environment is designed in order to allow end users achieve their goals only relying on their professional skill and knowledge. ICL is designed to solve the problem of tuning a system for compressing data still maintaining the information necessary for medical purposes. To this end, we adopt a) a *selective compression* technique [6, 7] which maintains the whole information in the regions of diagnostic interest (RoDI) and compresses the remaining parts of the image maintaining information sufficient to understand the clinical situation; b) a *visual interaction* technique [8] by which the radiologists can tune the control parameters avoiding the burden of mathematical details but experimentally observing the effects of their decisions.

The paper is organized as follows: section 2 contrasts the traditional scenario of the radiological lab with the new scenario of computer-based lab; section 3 describes the basic components by which ICL is built; in section 4 the image compression engine is introduced; in section 5 it is shown how the ICL is built from the basic components which have been specialized to image compression; last section draws conclusions.

2. Two scenarios

In a traditional environment, radiologists reach their diagnosis looking at radiological images and focusing their attention on *characteristic structures* (*cs*), i.e. image structures to which they associate a functional meaning or which constitute a perceptual unit for them. Often, they also emphasize 'important' css by specific marks, e.g. by hand-written circles. Radiologists exploit the full information of the important regions to reach their diagnosis, while use the css in the outside regions as a context supporting their reasoning. They often associate the marked css a text, describing the css meaning. In this way they create what has been recently defined a *characteristic pattern* (*cp*) [8]. Css are the visual clues, which allow the radiologists to identify a situation, and

0-7695-0040-4/99 $10.00 © 1999 IEEE

when associated to a meaning, i.e. used as a cp, allow them to accomplish their goal, for example derive a diagnosis or a prognosis. The marked image, the set of texts describing the css, the data which allow to put each css in relation with its description are called visual sentence (vs); and the archive of images associated to their interpretation (diagnosis) produced by the radiologist, being a set of vss, a Visual Language (VL).

When using a computer-based system to support radiologists in their activities, css and cps play important roles but the development and use of computer based system requires their precise definition. Cps are the goal of interactive or automatic image interpretation, indexes for image archiving and retrieval [9] and also the mean for communication between radiologists and system [8].

The description of the meaning associated to a cs must be expressed in a form interpretable by the system, typically as an attributed symbol [8]. The description of the meaning of the whole image i is in this case organized as a set d of attributed symbols, each one describing a cs. A cs in turn must be managed by the system as an atomic pictorial entity and also the relation between a cs and its description must be made explicit defining two functions: int, mapping the structure into the attributed symbol, i.e. a set of symbols, and mat, mapping the symbol into its structure. A cp is then described by a triple: cp=(<set of pixels>, <attributed symbol>, <int,mat>). The set of cps associated to an image express the interpretation of the radiologist and can be organized as a visual sentences (vs), a triple, vs=<i, d, <INT, MAT>> where i is an image, d is a description, INT a function mapping the css of i into symbols in d describing their meaning and MAT a function mapping symbols in d into css in i.

Note that in the Visual Sentence theory the image i is intended in a broad sense: an image is what can be displayed on a screen, i.e. texts, pictures, graphs. In a text the letters as well as words are at different level of abstraction the css of interest. Adopting this theory, we expect to develop an interactive system by which the radiologist can identify and mark the *Regions of Diagnostic Interest* (*RoDIs*) around the important css whose information must be fully preserved and decide how much information may be lost to compress the rest of the image. However, it is very difficult to formally assess the acceptable level of image quality. To overcome this problem, we adopt the informal, operative definition of *diagnostic quality* proposed in [6, 10]: "a lossy reconstructed radiological image can be accepted from a diagnostic quality point of view, if different radiologists with same qualification level can establish the same diagnosis both on the reconstructed image and on the original one". Hence, radiologists need a system, which support them in compressing the images, in judging the results and in repeating these activities until they are satisfied by the compressed image.

In this new scenario, a radiologist and the system should be able to perform the following activities:

a. the radiologist focus his/her visual attention on the image and detect structures of interest (css);
b. the radiologist interactively marks the RoDIs and the system captures this "selective" information (definition of RoDIs);
c. the radiologist interactively sets the compression parameters and the system captures them;
d. the system fires a compression engine using RoDIs and parameters to selectively compress the image;
e. the system fires a reconstruction engine, which presents the reconstructed image to the radiologist;
f. the radiologist judges the results; if the quality of the image is unacceptable according to his/her diagnostic quality criterion, the radiologist refuses the compressed image and goes back to c to change the setting of the compression parameters;
g. if the quality of the image is acceptable, the radiologist accepts the image, and the system stores the compressed image.

ICL is an environment developed to allow a radiologist to compress images following this procedure. This compression activity is seen as a particular case of image interpretation, in which the goals are a) the recognition of the regions of diagnostic interest, in which all the information conveyed by the original image must be preserved, b) the generation of an image, in which the different regions are selectively compressed according to their interest for the diagnosis, c) the generation of an image description - the interpretation of the image.

3. The basic components of a visual environment

ICL is a *Cooperative Visual Environment (CVE)* designed for image interactive compression and developed using the basic module Interactive Image. A CVE is a visual sentence formed by two parts: a *surface*, visible at any given moment to the user, and a *navigational and computational structure* (nc-structure), that captures the end-users actions, computes the reaction to them and represents the results of the computation on the surface- see Fig.1. The nc-structure is organized into three layers, and each layer is in turn a network of agents exchanging messages following standardized paths. At a bottom layer, a network of *executor agents* performs the computational activities. At a middle layer, *observer agents* monitor the activities of executors agent. Each observer is able to capture the current values of some variables from one or more executor, to receive data from other observers, and to organize them according to the rules of some pictorial language. Each observer is linked to one *surface agent* in the top layer. Each surface agent is responsible for mapping the results of the activity of one observer in well-determined region of the screen and to capture the events related to that screen region and feeding them to the observer: it adapts the input/output of an observer agent to the specific input/output devices [11]. Surface agents can communicate among them to synchronize their activities.

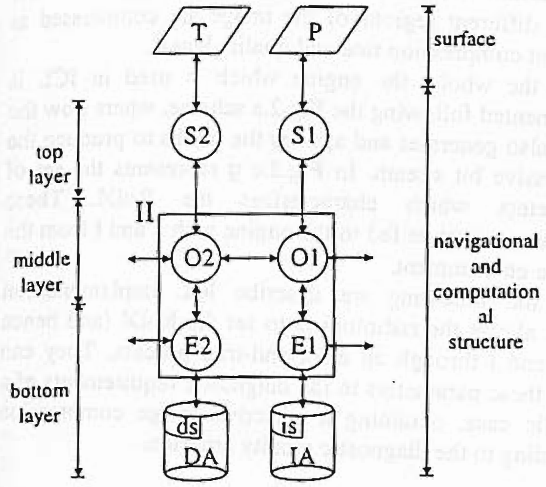


Fig.1 Architecture of an Interactive Image (P stands for pictorial data, T for textual data)

An *Interactive Image (II)* is a network of CVE agents composed of 2 executors (E1 and E2) and 2 observers (O1 and O2) (framed in Fig.1). The network is designed for image and situated description processing, creation and management according to the approach presented in [3]. An II can interact with image archives (IA) and description archives (DA) and with surfaces elements (S1 and S2). IA stores images and image schemes (is), which specifies how images are organized and the rules for their management. DA stores the descriptions and the description schema ds, which specifies how descriptions are organized and the rules for their management.

In II, E1 is able a) to read is, b) to read image data from the image archive (IA), c) operate on image data, d) submit its pictorial results in a standard format to O1, e) store image data in IA. E1 may also send/receive synchronizing messages from executors outside II.

O1 is an observer, able to a) capture a set of images data from E1, b) receive image data from other sources, c) merge the available image data, d) prepare the image data to be fed to a surface agent (S2 in Fig.1).

E2 is an interpreter, which can a) read a description schema form the archive DA, b) interpret image data according to the rules stored in a description schema, c) manage a description according to the rules stored in a description schema, d) store a description in DA.

O2 is an observer, which exchanges data with E2, and prepares them for representation. If data are to be represented as images, O2 prepares and feeds them to O1; if data are to be represented in textual form, O2 prepares and outputs them to surface agent (S2). It also receives data from its surface agent.

In the following we will use II as a basic module to implement ICL. Our aim is to build a tool which allows radiologists to achieve a spatially inhomogeneous distribution of reconstruction quality in that the regions of the image which are of their interest are compressed without any loss of quality, while the context is compressed according to the diagnostic quality request introduced in sec.2. This tool uses as an executor an image compression engine, whose control parameters are interactively defined by a radiologist, which judges the quality of reconstructed images by visual feedback.

4. The image compression engine.

To implement the image compression engine, we adopt a recently introduced refinement of a compression scheme based on the Discrete Wavelet Transform (DWT) [12], on the zerotree quantization [13, 14] and on the arithmetic codification of the coefficient. Fig.2.b describes the classical ZeroTree Coding (ZTC) scheme. Here DWT maps an input image into a set of wavelet coefficients organized in 2 dimensional multiresolution data structure. Then the ZeroTree Quantifier (ZTQ) maps this structure into a progressive string of symbols, each symbol describing the importance of the contribution of a wavelet coefficient to the reconstruction of the image. Here a tree is a set of DWT coefficients sharing the same spatial localization at different resolution level; a ZeroTree symbol codifies a tree of coefficients when all of them are lower than a certain threshold T. Due to the properties of DWT, the ZeroTree symbol is generated with a high probability; this allows a prediction of many coefficients with minimal coding effort. This coder iteratively scans the coefficients, halving T at each iteration, performing a bit-plane coding and producing a "progressive" refinement of the quantization level.

Last, Arithmetic Coder (AC) codifies the progressive string of symbols into a progressive bit-stream without loosing any information. In principle this procedure can code the image into a progressive bit stream without any loss of information, but reaching a low compression rate.

To improve the compression rate two strategies are available which, however, are not lossless. The first strategy establishes a target quality q which is an input to ZTQ. ZTQ builds a reduced string of symbols from which an image of the desired quality can be reconstructed. The second strategy consists in cutting the bit stream according to a rate criteria r, which is an input to AC. Fig.2.a shows the scheme of an image compression - decompression engine using a progressive bit stream. In the case of medical images the second strategy cannot be adopted because does not allow any control on the quality of the reconstructed image.

However, also the first strategy has its weakness from the point of view of the radiologist, because the quality is a global characteristic of the image established according to objective statistical properties of the image itself. In this way it may occur that the regions of interest for the radiologist can be degraded even if the quality of the whole image remains in the range considered acceptable. To avoid this fault a new method has been proposed in [6, 10], which imposes a spatial hierarchy of importance on the wavelet coefficients to preserve the quality of the image in the regions of interest; in the remaining part of the image, the quality may be reduced as far as the radiologist is able to understand the context. This hierarchy is imposed using a mask (Fig.2.c) which weights each coefficient according to its importance in RoDI reconstruction.

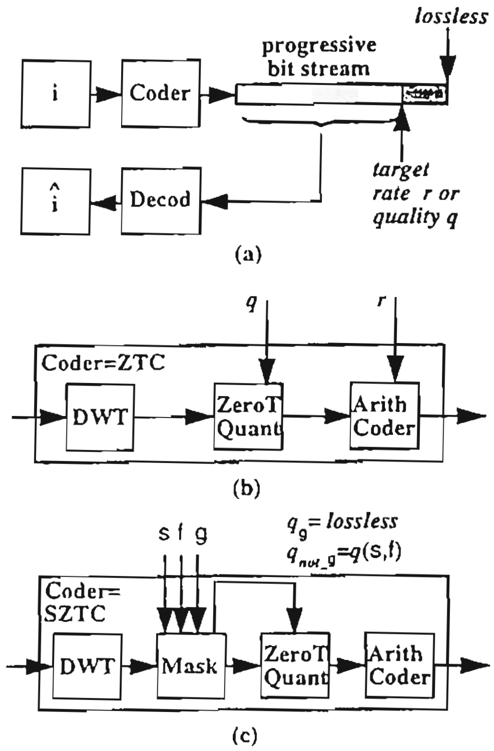


Fig.2. a) the general architecture of the image compression engine; b) the classical scheme for ZT coding; c) the scheme for selective ZT.

More precisely, by the use of the masks, the wavelet coefficients belonging to the RoDI are multiplied by a parameter s (salience), which express the relative importance of the RoDI with respect to the context. If s rises, the rate of compression rises, but the context, i.e. the part of the image outside the RoDI, becomes more and more blurred and some artifacts may arise. A second parameter f (focalisation) determines the smoothness of the quality drop between RoDI and context. If f rises, the mask will result steeper, and the compression rate higher, but the transition from RoDI and low quality context sharper. s and f are used to automatically compute the target quality parameter, that is in turn used by ZTQ to stop the compression. The whole activity results in an image in which quality outside RoDI is adequately degraded, while RoDI compression is lossless. This compression is called Selective ZeroTree Coder (SZTC), in that different regions of the image are compressed at different compression rate and quality levels.

On the whole, the engine which is used in ICL is implemented following the Fig.2.a scheme, where now the coder also generates and applies the masks to produce the progressive bit stream. In Fig.2.c g represents the set of parameters which characterizes the RoDI. These parameters must be fed to the engine with s and f from the outside environment.

In the following we describe ICL implementation which allows the radiologists to set the RoDI (and hence g), s and f through an error-and-trial process. They can adapt these parameters to the diagnostic requirements of a specific case, obtaining a selective image compression according to the diagnostic quality criterion.

5. An overview of ICL.

In this section we describe how to merge the CVB technique and the discrete wavelet transform to implement ICL in an interactive system usable by radiologists. ICL is organized as a(n electronic) logbook as defined in [5]. "A logbook is a vs, whose surface image is divided in pages which can be presented on the screen one at a time. ICL is a logbook constituted of three pages. The first ICL page is dedicated to the image selection and loading from IA: its structure and functionality will not be discussed here. The second page is dedicated to the RoDI definition and qualification, while, the third page allows the visual comparison between the original image and the compressed image, generated by the compression engine, according to the parameters established by the radiologist interacting with the second page.

Fig. 3 shows ICL second page with the part of the nc-structure generating and managing it. This second page is composed of four different windows. The first window named 'Image Compression Laboratory', is common to every page and represents the logbook cover, which gives the radiologists some general information on the environment they are using and whose nc-structure will not be discussed here.

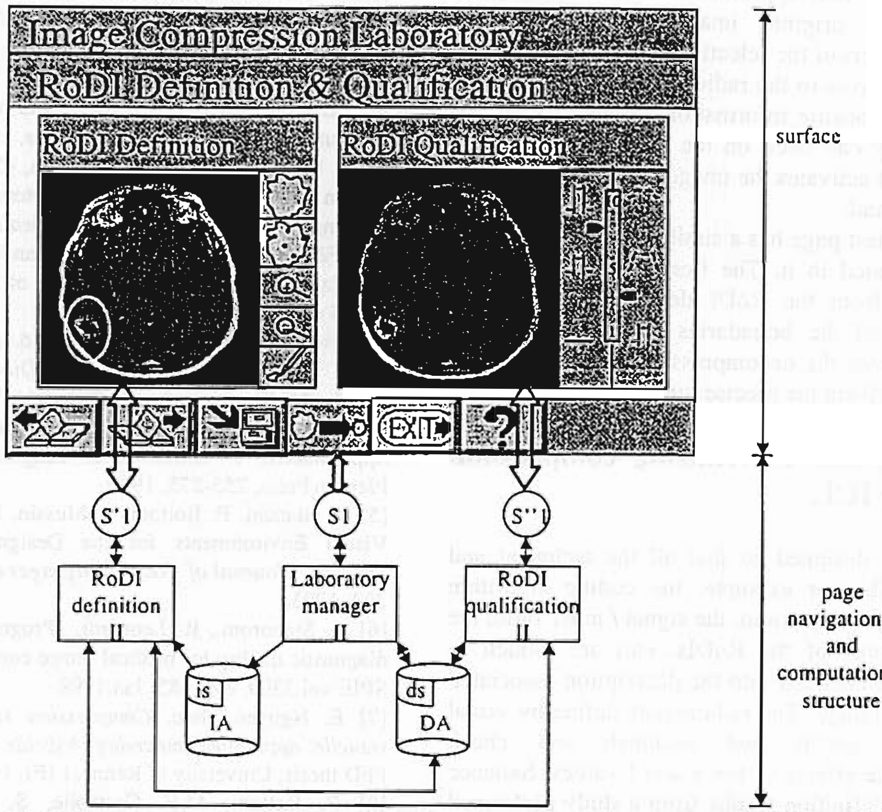


Fig. 3. The RoDI definition and qualification page and the underlying II architecture

The second window is contained in the first one: it appears below the logbook title and displays the page 'RoDI Definition and Qualification' and shown by its title. Interacting with this page the radiologist can perform one out of six different actions at a time. These actions are denoted by the six icons displayed on the bottom of the page. The first (second) icon denotes the action of going back (going forward) to the preceding (following) page. When one of these icons is clicked on, the selected page appears on the screen. The logbook cover remains unchanged, but the window contained in it is fully updated.

The third icon denotes the activity of storing in IA and DA the results obtained so far -new or updated descriptions and the new or updated images. The fourth icon fires the compression engine; the fifth closes the interactive session and the last one activates the help program.

This page window hosts two other non-overlapping windows, which allow the performance of two different activities as denoted by their titles. The first one allows the 'RoDI Definition' on the original image; the second one allows the setting of the two parameters s and f.

The second page nc-structure is a network of three interactive images, each one connected to only one surface agents, because each one only manages only one window (surface element). Every interactive image can read/store images and descriptions from/into IA and DA.

The RoDI definition II, autonomously presents in its window the image selected by the radiologist using the first page. This II allows the execution of five actions. The first action is denoted by the first icon on the top of the menu on the right of the image, and allows the representation of the original image, hiding the RoDI contours if defined. The second action is denoted by the second icon, which shows the contours of the defined RoDI over the original image. The two next actions, denoted by the two next icons, allow to zoom in and out the image. The last action, denoted by the writing-pen icon, allows to define the contour of the RoDI. RoDI are here framed by ellipses generated by the following procedure: a) click on the action icon, b) move the pointer to the position of interest, by moving the mouse, c) press the mouse button, d) move the pointer maintaining the button pressed until the ellipses reaches the desired dimensions, e) release the mouse button. As a reaction to this procedure, the RoDI definition II generates the ellipses on the desired position and after gesture e) stores in DA the description of the ellipses itself.

The RoDI qualification II, autonomously presents in its window the same image appearing on the 'RoDI Definition' window. After each RoDI is defined, RoDI qualification II generates the colored cloud simulating the effects of the selective coding with the values of s and f pointed at by the two sliders in the image. If the radiologist is not satisfied, he/she can change one or both the values of s and f by moving two slide-controls. The vs reacts proposing a first approximate visual feedback: it superimposes the original image a colored cloud simulating the effects of the selective coding. This is a fast operation, which gives to the radiologists a first feedback on the effects of loosing information. When radiologists are satisfied, they can click on the fourth button in the page menu, which activates the image compression engine on the image at hand.

The third and last page has a similar structure with two windows represented in it. The first window shows the image resulting from the RoDI definition, that is the original image and the boundaries of the RoDIs. The second image shows the decompressed image and allows radiologists to perform the precise tuning of s and f.

6. Evaluating and customizing compression techniques by ICL.

ICL has been designed so that all the technical and algorithmic details (for example, the coding algorithm adopted, the compression ratio, the signal / noise ratio, the geometrical features of the RoDIs, etc) are hidden to radiologists and embedded into the description associated with each active image. The radiologists define by visual interaction the RoDIs and establish and check experimentally the effects of the s and f values. Salience and focalisation definition results from a study performed by one of the authors [6, 10] in collaboration with the radiologists of the Neuroradiology Department of the University of Brescia. This experience also supported the proposal of the quick feedback in RoDI qualification.

ICL is now under development, and usability experiments are foreseen to confirm the usability hypotheses, which underlie its design. These experiments will answer the question if the proposed VL is the better customization for radiologists. A second set of experiments is foreseen, that is the evaluation of the proposed compression technique against the traditional ones. The modular architecture of ICL will allow its adaptation for the execution of the two types of experiments, in the first case by adapting the surface, in the second case by substituting the compression engine in the Laboratory Manager. These experiments will be the next steps in our research.

ACKNOWLEDGMENTS.

We would like to thank Dr. Roberto Gasparotti, for fruitful discussions regarding the overall methodology and for a direct participation as an expert radiologist during the experimental simulation of this work, and the Neuroradiology Department of University of Brescia.

REFERENCES.

[1] J.C. Brancheau, C.V. Brown, "The management of end-user computing", *ACM Computing Surveys*, 25, 437-482, 1993.

[2] P. Bottoni, U. Cugini, P. Mussio, C. Papetti, M. Protti, "A system for form-feature-based interpretation of technical drawings", *Machine Vision and Applications*, 8, 326-335, 1995.

[3] F&DA Center for Devices and Radiological Health, "Guidance for the content and review of 510(K) notifications for PACS and related devices", *http://www.fda.gov/cdrh/ode/odeot416.html*, 1996

[4] E. Barichella, M. Beretta, N. Dioguardi, P. Mussio, M. Padula, M. Pietrogrande, M. Protti, "A Visual Environment for Liver Simulation Studies", in *Visual Languages and Applications*, T. Ichikawa, E. Jungert, R. Korfhage (eds.), Plenum Press, 255-275, 1990.

[5] N. Bianchi, P. Bottoni, P. Mussio, M. Protti, "Cooperative Visual Environments for the Design of Effective Visual Systems", *Journal of Visual Languages and Computing*, 4, 357-382, 1993.

[6] A. Signoroni., R. Leonardi, "Progressive ROI coding and diagnostic quality for medical image compression", in *VCIP'98*, SPIE vol.3309, 674-685, Jan.1998.

[7] E. Nguyen Phuc, *Compression sélective et focalisation visuelle: application au codage hybride de séquences d'images*, PhD thesis, University of Rennes1 (F), 1995.

[8] P. Bottoni, M.F. Costabile, S. Levialdi, P. Mussio, "Specifying dialog control in Visual Interactive Systems", *JVLC*, vol.9, n.5, pp.535-564, 1998.

[9] P. Bottoni, P. Mussio, M. Protti, "Metareasoning in the determination of image interpretation strategies", *Pattern Recognition Letters*, 15,177-190, 1993.

[10] A. Signoroni., R. Leonardi, "Progressive medical image compression using a diagnostic quality measure on regions-of-interest" *EUSIPCO'98*, 2325-2328, Sep.1998.

[11] P. Bottoni, P. Mussio, B. Olivieri, M. Protti, "A completely visual environment for agent-based computing", *Proc. AVI'98*, T.Catarci, M.F.Costabile, G.Santucci, L.Tarantino eds., ACM Press, 261-263, 1998.

[12] M. Vetterli, J. Kovacevic, *Wavelets and Subband Coding*, Prentice Hall PTR, 1995.

[13] J.M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients", *IEEE Trans.on Signal Proc.*, 41(12), 3445-3462, Dec. 1993.

[14] A. Said, W.A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees", *IEEE TCSVT*, 6(3), 243-250, Jun.1996.