Image Analysis and Computer Vision in Medicine

Thierry Pun[1], Guido Gerig[2] and Osman Ratib[3]

[1]Artificial Intelligence and Vision Group
Computing Science Center, University of Geneva
24 rue Général Dufour, CH-1211 Genève 4, Switzerland
e-mail: pun@cui.unige.ch

[2]Image Science Group
Communication Technology Laboratory
Swiss Federal Institute of Technology in Zürich (ETHZ)
ETH Zentrum, CH - 8092 Zürich, Switzerland

[3]Digital Imaging Unit
University Hospital
24, rue Micheli-du-Crest, CH-1211 Genève 4, Switzerland

Corresponding author: T. Pun.

Submitted to: Computerized Medical Imaging and Graphics, special issue on Multimedia techniques in the medical imaging environment.

Abstract

Multimedia lives with images; medical images are born from digital imaging. A physician's multimedia workstation cannot exist without tools for manipulating images, performing measurements and, generally speaking, extracting and collecting pieces of information from the available data.

Image analysis and computer vision constitute a wide and rapidly evolving field. This paper is intended as an introductory document for medical imaging researchers and practitioners wishing an overview of recent developments and trends. The major lines of activities in the domain are presented, under the common framework of a processing pipeline. After a presentation of the various stages of this pipeline, current subjects of research are indicated.

Keywords

Medical imaging, image analysis, computer vision, features extraction, segmentation, reconstruction, matching, recognition, artificial intelligence.

1. Introduction

1.1 Digital Imaging

A *digital image* is characterized by its lattice and by the corresponding picture elements (1). The *lattice* is typically a rectangular grid of dimension 2 or 3, corresponding either to 2 or 3 spatial coordinates (planar or volumetric data), or to 2 coordinates plus time (images sequence); in general, the dimension of the lattice is n, with $n \geq 2$. In the 2D case, individual elements may have four or six nearest neighbours, corresponding respectively to rectangular or hexagonal arrangements. In the 3D case the lattice is usually parallelepipedic, with six neighbours. The values of those picture elements, called *pixels* in 2D and *voxels* in 3D, can be single data such as a density, pairs of data such as relaxation times $\{\tau_1, \tau_2\}$ in MRI, triples such as trichromatic values, etc.

Image manipulation encompasses several types of processing techniques, collectively known as *digital imaging*. *Image synthesis* (or *computer graphics*) consists in assigning a value for each of the above pixels or voxel, the end product being a 2D or 3D image with (generally) some semantic and/or pictorial content. Very generally speaking, image synthesis is a parameters to image transformation (2). These parameters can be numeric functions, as in *scientific visualization*, or, in the medical imaging context, they can be a set of anatomical objects stored as surface or volumetric models or as a parametric description. Image synthesis has always been closely associated with human-computer interaction. Most reference textbooks and publications describe interactive computer graphics and not just image synthesis per se. The ability to change pixels rapidly leads naturally to interactive generation of static and dynamic pictures. The full exploitation of the computer as a medium for image synthesis is thus only achieved in a dynamic, interactive environment.

Conversely, image analysis consists of a picture *to parameters* transformation, starting with image(s) for which each element has a known value. It is customary to subdivide this domain into *image processing* (1), *image analysis* (1,3), and *computer vision* (or *image understanding, scene analysis*) (3,4,5). Image processing consists in transforming one image into another image, often with the same support; the purpose is typically to eliminate or enhance some features. Image analysis goes one step further by extracting parameters and analyzing them; classical domains of application are medical imaging or industrial robotics. Finally, computer vision is a field of study which aims at designing computer systems mimicking the human sense of sight.

A more restricted goal for computer vision is to provide a symbolic description of an image or scene. Computer vision methods usually encompass both numeric and symbolic processing of optical data, which is for exemple obtained from multi-camera input. By extension, “computer vision” is often used in place of “image analysis” and even “image processing”, as soon as the algorithms involved exceeds a certain level of complexity. It could be argued that medical imaging relates more to image analysis than to computer vision; recent devel-

opments in medical imaging show however that this distinction is becoming thin.

1.2 Image Analysis and Computer Vision in Medicine

There are various motivations for using digital image processing methods in medicine:

- *new modalities and multimodal analysis*: foremost comes certainly the possibility of exploring new imaging modalities, leading to new anatomical or functional insights; further, image analysis will support the combined evaluation of data from different modalities;
- *morphometry*: the use of computerized techniques allows better precision and repeatability, with, as a consequence, improved objectivity of measurement of morphometric parameters like size, area, volume, circumference, etc.;
- *improved interpretation*: the sensitiveness of those new imaging modalities coupled with the power of recent visualization techniques enable more refined diagnosis than with using conventional exploratory methods;
- *more accurate prediction*: a consequence is the ability of providing more finely tuned medical treatment, for example lower doses in radiation therapy or more accurate positioning in head surgery;
- *process automation*: many medical operations can benefit from the reliability provided by automatic processing, from the screening of biological specimen to vision guided surgery;
- *understanding of volume data*: recognition of structures from volume data is not a spontaneous visual task and will benefit from computerized processing and visualisation.

Medical applications of image synthesis techniques are mostly for 2D and 3D visualization purposes (6,7). Typical examples are in diagnosis or planning, for example for surgery or radiotherapy. Graphical methods can also be employed for simulation, typically by means of computer animation techniques (8). It is worth noting that other theoretical concepts which are usually perceived as pertaining to computer graphics play an important role in computer vision; this is discussed in §2.2. The links between these two kindred fields are numerous; it is hard to conceive a physician's workstation without methods originating from both.

The history of image analysis in medicine is older than the one of image synthesis (1960's versus than 1980's) (9,10). The early works were concerned with applications such as chromosome pairing, morphological classification and counting of particles (11), coding for image compression (12), storage and communication (13,14). More recently, the advent of new imaging technologies has lead to methods exploring 3D structures and dynamic changes of objects (15,16,17), usually starting from slice data or possibly from time sequences of 2D or 3D images. Extraction of 3D morphological structures is a very active research area

(e.g. (18,19)), possibly using artificial intelligence methods (e.g. (20)). Complex computer vision techniques are also being used in domains such as medical robotics (21). Although 2D analysis still receives considerable interest, the trend is now to directly analyse volume data by 3D methods.

1.3 Outline

This paper is intended as an overview of recent developments and trends. It wishes to provide a short survey of the principal families of approaches, their use and limitations. Algorithms generic or specific to particular imaging modalities are not given here (see for example (10)).

Section 2 presents the analysis pipeline, which accepts as input raw data and provides some form of processing or interpretation. Two basic processing paradigms are summarized in section 3, namely linear (or pseudo-linear) and morphological. Section 4 develops the problem of features extraction by means of image segmentation techniques. Analysis of multi-slice and volume data, and matching, between data of identical dimension, are respectively presented in sections 5 and 6. Common approaches for identifying objects are summarized in section 7. What the authors perceive as important trends are indicated in section 8.

2. The Analysis Pipeline

2.1 Processing Steps

Medical imaging devices usually provide 2D data (pixels); more recent sensors now allow direct acquisition of 3D data (voxels). The data can then either be visualized, using 3D image synthesis methods, or analysed, using 2D or 3D image analysis and computer vision techniques. The processing steps that are required to perform an analysis are applied in sequence, hence the idea of an *analysis pipeline* through which the input data is circulated†.

The generic image analysis and computer vision pipeline is composed of the following steps (Fig.1):

- image acquisition, from sensor to digital image: this first stage includes all processings that are required to create the intensity data (e.g. reconstruction from projections

†. Conversely, synthesizing an image can also be done by means of a *visualization pipeline*, composed of the following steps: modelling three-dimensional objects in their own coordinates system, composing a scene by instancing these objects at various locations within a common reference frame, lighting and shading the scene, projecting visible parts onto the display.

if needed);

- preprocessing: this step includes the data manipulations that are performed either systematically on all images (e.g. for noise reduction), or interactively by an observer (e.g. for enhancement);
- segmentation: the purpose of this processing stage is to subdivide the image into its constitutive parts, such as regions and edges, in order to obtain significant features;
- reconstruction of data: series of m-dimensional data are matched in order to provide n-dimensional information ($m < n$) (the visualization itself by 3D rendering is not included here). The classical example consists of matching information from 2D slices in order to obtain a 3D volumetric representation;
- matching: series of n-dimensional data are put into registration in order to combine information. This can be applied to data from different sources, or from the same source at different times;
- recognition: identification of elements present in the scene, which is often but not always the final purpose of the analysis.

Figure 1 here

As shown in Fig.1, variations from this unidirectional flow of information are possible. For example, after an nD reconstruction, it is again possible to apply segmentation on the nD image. Also, after having recognized part of (or all) the elements in the image, it is possible to reiterate the whole processing sequence with the additional knowledge provided by the information already extracted. Although not yet common in medical imaging, the usage of such *top-down information* is increasingly encountered in “theoretical” or industrial computer vision (20,22).

According to this rather generic pipeline, some typical medical imaging applications can be decomposed as follows:

- enhancement (e.g. for contrast manipulation, pseudo-colouring):

  `preprocessing → 2D or 3D visualization by image synthesis;`
- morphometry:

  `preprocessing → features extraction by segmentation → morphometric measurements;`
- classification:

  `preprocessing → features extraction→ recognition by means of morphological parameters;`
- 3D imaging and diagnosis (e.g. for radiotherapy, surgery):

  `2D or 3D acquisition → segmentation → approximation → generation of 3D model → visualization;`

- multimodality (functional) imaging:

  `mD acquisition` → `segmentation` → `approximation` → `nD reconstruction` → `matching` → `visualization;`
- data coding and storage (e.g. for PACS):

  `2D or 3D acquisition` → `(preprocessing)` → `approximation.`

These various steps are presented in more detail in the following sections.

2.2 Relationships with Image Synthesis

Although this article concentrates on image analysis and computer vision, image synthesis is in many respects linked to these disciplines (23,24). The most immediate example is the data visualization itself. The physician's workstation will foremost be a display device, showing the results from the reconstruction and/or analysis. In addition to this obvious example, various concepts "borrowed" from computer graphics are often employed in scene analysis. The opposite is of course true, but of less interest in the context of this paper.

The common basis for image synthesis and analysis lies in the way humans visually perceive things. Computer graphics is obviously intimately concerned with perception. The physics of display certainly has influence over what is observed, but what really matters is the internal representation to which a given image refers to and which it seeks to recreate in the user. In the case of a physician establishing a diagnosis on the basis of a 3D rendered image, using knowledge from human perception could help design analysis and synthesis algorithms that would convey a message more forcefully.

From a more pragmatic standpoint, graphics concepts that are important in image analysis include light models for illumination and reflection, geometrical transformations, contours and surface approximation, algorithms for region delineation, and, of particular importance in 3D medical imaging, geometrical modelling.

Another joint area of interest is in hardware and software, with a common trend towards the use of RISC processors, parallel computation, color devices, etc. Mass storage using jukeboxes of optical discs becomes common, especially in view of *Picture Archiving and Coding Systems* (PACS). Various software standards are also emerging, for image synthesis (PHIGS+, Programmer's hierarchical interactive graphics system), for interface design (OSF/Motif), for communication (X11, PEX), for image storage (e.g. ACR-NEMA specifications for medical images), for image compression (JPEG, MPEG).

3. Basic Image Processing Paradigms

Five families of methods are often said to be basic image analysis paradigms: linear or pseu-

do-linear (1), morphological (25), statistical (26), syntactical (3,27), neuronal (28). Only the first two are presented below, since the last three families of methods mostly belong to one particular stage of the pipeline, namely object recognition. Furthermore, neural networks are non-linear classifiers (29), which would tend to make them part of the statistical approaches.

3.1 Linear and Pseudo-Linear Paradigms

Linear systems theory is the basic paradigm for a large number of methods (1). Having an nD discrete input image I_x, the effect of applying a discrete linear operator (filter) of known impulse response (transfer function) g yields $I_y = I_x * g$, where $*$ is the *discrete convolution* product. Although not many algorithms are linear, it is often useful to approximate and study their behaviour using the linear formalism.

Fourier transforms are used for spectrally characterizing data or (linear) operators, for designing filters, for analyzing periodical structures, for sampling and interpolation problems. Also, it is at the basis of the projection methods for reconstructing slices (Fourier slice theorem and backprojection).

Correlation, which determines the degree of correspondence between a pattern and an image, is used either for isolating a priori known shapes, or for alignment purposes (30). Depending on the size of the pattern, the matching operation is implemented either in the image or in the transform domain. Similarly, the well known Hough transform for detecting straight lines or circles can also be considered as belonging to the family of pattern matching methods (31).

3.2 Morphological Paradigm

An alternative to linear methods is *mathematical morphology*, where set operators allow a more formalized mathematical treatment of many image analysis problems (25,32). The basic idea of mathematical morphology is akin to pattern matching: a particular pattern called *structuring element* is positioned on every image pixel. At each position, a logical test is performed in order to check whether the image is locally identical or not to this structuring element. Depending on the nature of the test, various operations are accomplished.

The basic operations, the most commonly used also, are dilation, erosion, opening (erosion followed by dilation) which eliminates small objects, and closing (dilation then erosion) which fills holes and joins elements. Built upon those elementary operations, a whole family of morphological operations allow applications such as noise cleaning, particles singulation, counting, morphological parameters extraction.

Despite its potential power, mathematical morphology is more frequently used in biomedical image analysis than in “true” medical imaging. One of the reasons for this situation is that the theoretical framework has been essentially developed for binary 2D pictures. And

despite the attention that grey-level morphology is gaining (32), applications of mathematical morphology to images of dimension greater than two are yet to come. Also, the alphabet of morphological operations is in theory defined on a hexagonal lattice, which might require transforming the original, rectangularly arranged data; camera with hexagonal sensor grids are still uncommon.

4. Features, Preprocessing, and Segmentation

Preprocessing aims at eliminating systematic perturbations from the image, or at enhancing the image for visualization and possible further processing (1). Typical perturbations are high-frequency acquisition noise, background luminance variations, camera geometrical distorsions, echo, etc. Preprocessing methods are often applied in a systematic fashion on all images created by a given device. They are therefore often device-dependant; also, they need to be fast and efficient.

Segmentation aims at extracting significant primitives and regions of interest (ROI). This is necessary either for performing measurements, that will allow classification and identification, or for reconstruction, in order to determine the elements that will be put into correspondence for creating a 3D representation. An enormous corpus of methods has been published, with three families of approaches being customarily defined: thresholding, edges extraction, regions extraction; the general principles indicated below apply to both 2D and 3D images.

4.1 Measures and models

The main function of preprocessing and segmentation is to extract *measures* for solving classification and identification problems. Although there are as many different measures as there are particular problems, certain basic descriptors are very often used (1,3). In two-dimensions, contour-based measures are typically the length L, the average contrast on each side of the contour, the local or average curvature. Region-based descriptors are usually based on simple values such as average luminance or color, surface S, perimeter p, compactness p^2/S, average luminance gradient with neighbouring regions. Moments about the center of mass are also used, for example for determining the axes of an approximating ellipse; the position of this center can help to locate objects. These measures can easily be generalized to higher dimensions.

Segmentation is also necessary for extracting primitive elements that will then be approximated by some parametrized shape, as well as for isolating regions or contours in interactive image manipulation and observation. Those parametrized shapes, or *models*, play a fundamental role in image synthesis as well as in computer vision (23,33).

Geometric models are thus the basic tokens that the visualization pipeline will have to process for synthesizing an image. They are also essential in model-based segmentation, where the image is decomposed into parts that fit to these models. In the context of medical imaging, the most commonly used models are the simple voxel and the 3D polygonal planar patch. Non-planar models such as spline patches (2) and hyperquadrics (34) are sometimes used, but the trend is towards deformable models such as the 2D “snakes” (35) or 3D “inflatable balloons” (36) that modify themselves in order to adapt to some key data features. They not only allow a better fit to the data, but also inherently provide segmented regions.

4.2 Preprocessing

Photogrammetric methods are used when the artifacts have known spatial or luminance characteristics; examples are lookup tables for linearity correction, dynamic range modifications, histogram equalization or interactive histogram modifications for visualization, local or global artifact estimation and subtraction. They are often performed in real time in the display memory (37), in which case they mostly serve a visualization purpose.

Filtering methods are preferred when the spectral characteristics of the artifacts are known. Low-pass filters (average, Gaussian) are used for decreasing noise, band-pass filters to eliminate periodical perturbations, high-pass filters to enhance and sharpen edges (*unsharp masking*). Non-linear operators are also employed, such as the median or out-of-range filters. Filtering is often performed off-line or sometimes using specialized circuits.

A particular class of algorithms (related to low-pass filtering) is used when an enhanced version of one object is desired, assuming that many degraded, translated and rotated versions of the same object are available. The basic idea of the method, called *correlation averaging*, is to align the elements by correlation techniques and then average them (30). The signal-to-noise ratio, hence the spatial resolution, roughly increases as the square root of the number of averaged individual elements.

Geometric corrections are required when images are geometrically distorted. This is for example necessary for registering images from various sources. Finally, in this taxonomy of preprocessing families, *restoration methods* are used when a model of the perturbation allows the derivation of an optimal cleaning filter.

In summary, for visualizing data on a medical imaging workstation, look-up table manipulations and unsharp masking are often helpful; in this case, care should be exercised in order not to have the preprocessing modify the picture in such a way as to mislead the observer and generate wrong interpretations. For preprocessing data before further image analysis, low-pass filtering is usually necessary.

4.3 Segmentation

Thresholding methods provide means for determining one (or more) grey-level thresholds,

usually by analyzing the smoothed global gray-level histogram of the image. The result is thus a binary image composed of regions, hopefully with objects segregated from the background. The major drawback is the lack of local adaptability: it is rarely the case that global thresholding provides satisfactory regions. It is in general necessary to post-process the image, typically using mathematical morphology algorithms such as erosion/dilation.

Edge extraction allows delineation of regions by locating their contours. These transitions, that is areas of high variation, are initially extracted using first derivative operators (mostly Roberts' or Sobel's), second derivative operators (Laplacian mask), or adapted operators that locate which parts of the image match a model of an idealized contour (Mero-Vassy's or the trendy Canny-Deriche's for which fast recursive algorithms are available (38)). Final contours are then obtained by locating the gradient maxima in the resulting image by means of a peak-following algorithm, or the zero values of the Laplacian. Despite the apparent simplicity of the problem, extracting edges is still an area of research. Amongst recent trends are the use of pyramids, higher order derivatives, linear scale-spaces, linear combinations of directional derivatives, coherency detection using cooperation-competition or networks of oscillators, etc. (39). The major drawback of edge extraction for delineating regions is a high sensitivity to noise; typical post-processing include thinning algorithms and gap-closing methods (see (40) for a comparison and discussion of several edge finders). Pre-processing for noise reduction is also mandatory; it is often performed by means of a linear Gaussian or a non-linear median filter.

Region extraction yields parts of the image that satisfy a given uniformity criterion (1). Typically, it is required that for a given region the grey-level difference between a given pixel and any other one in the region be smaller than a certain value; a variation is to consider a given pixel and only its close neighbours. By principle, those methods provide closed regions and therefore closed contours, easily amenable to morphological measurements. The drawbacks are complexity of implementation, and the fact that usually many small regions are obtained. Here also post-processing is required, such as regularization of the shapes and elimination of small regions by erosion/dilation. In addition, segmentation of textured images is still a major problem (41); the current trend in this case is towards the use of compact support functions such as Gabor filters (42) or wavelets (43).

In summary, it is difficult to select an appropriate segmentation method. There is no "best approach"; interactive image manipulation softwares are therefore necessary (e.g. (44,45,46,47)). The least commitment principle should be observed, that is to avoid performing drastic operations whenever possible. As an example, if edges are desired, more information is preserved by doing edge extraction and peak following, than by thresholding the image and performing morphological erosion.

4.4 3D Segmentation

The third dimension brings additional information; it is therefore a good strategy to try to

use it as much as possible. Although it is feasible to reduce 3D edges extraction to a 2D algorithm operating on slices (38), the trend is to directly operate on 3D images (48).

Segmenting 3D images into regions of interest (or, generally speaking, nD images into nD regions) is somewhat similar to the segmentation of 2D images into 2D regions. The same families of methods can be used (49): thresholding, determination of boundary surfaces where the data varies, extraction of volumetric regions where a uniformity criterion is satisfied.

Results obtained using these classical approaches are often not satisfactory, basically because of the poor quality of the data. Other approximation principles are the subject of current research; a typical example is the use of 3D “inflatable balloons” (36) that adapt to key data features, such as discontinuity surfaces. This adaptation is accomplished by iteratively changing the model in order to yield an equilibrium state between internal forces (given by the elastic properties of the balloon) and external forces (attraction towards the key features).

5. 3D models from volume-covering acquisition techniques

The technological step from 2-D projection of 3-D bodies by transmission techniques (X-ray) to full 3-D imaging techniques (CT and MRI, e.g.) can be considered as a major breakthrough in medical imagery. This capability allows physical measurements to be obtained in the same number of spatial dimensions as the original human body. The possibility of visualizing 3D structures, both anatomical and functional, has brought new and extremely useful tools for diagnosis and treatment planning. 3-D techniques acquire the spatial information in the form of densely sampled volume elements (voxels), most often in a multislice process with coarser sampling between the slices (in order to avoid excess radiation or due to time limitations). If the inter-slice distance becomes comparable to the voxel dimensions within the slices, however, one can think of a contiguously sampled data volume.

The two different ways of considering the data, either as a spatial sequence of 2-D images or as a data volume, also influences the type of computer assisted analysis. The classical example consists of matching information from 2-D slices in order to obtain a 3D volumetric representation. The single slices of a sequence are processed by techniques developed for 2-D image segmentation, often not considering the context information available from consecutive slices. Regarding image data as a contiguous volume, however, leads to the development of true 3-D image processing methods, which most often represent extensions of 2-D algorithms.

The two major families of methods for manipulating 3D data are the voxel and surface approaches (50,51). With the *voxel methods*, voxels represent the basic elements of the data

structure. Processing such information is cumbersome in view of the large quantity of data. In addition, visualization is computationally complex: either the list of all voxels has to be inspected, with many unnecessary operations because of hidden voxels, or the surface of the set of voxels has to be determined (52,53). Alternative visualization methods for voxels do exist, and their performance has become absolutely comparable to other techniques rendering only the object surfaces (54). Voxel methods have the advantage, that the original image intensities can be easily combined with surfaces (e.g. on cutting planes) and that surgical planning is facilitated by manipulating volumes rather than hollow surfaces.

With *surface methods*, 3D information is often indirectly obtained from the 2D slices, by extracting and matching 2D primitives from these slices in order to build a 3D triangulation of (triangular) planar patches (55). The triangulation is typically based on the registration of polygonal contours approximations from the successive slices. Various conditions might be imposed, such as external surface maximization, or internal volume minimization. Although this approximation process might be difficult, it allows an important reduction of the amount of data, additional possibilities such as volumetric and mass measurements, and efficient visualization and animation schemes. It can also be useful as a first step before applying matching algorithms (§6).

After generation of 3-D object models, the result is usually visualized using image synthesis techniques. For anatomical images (e.g. CT, MRI), shading and sophisticated lighting can be used, whereas for functional images (e.g. PET), the original intensities should not be modified since they are the desired information; therefore, no shading or lighting algorithm should be applied. Functional information is often combined with morphological information by projecting the original intensities onto anatomically meaningful surfaces.

6. Matching

6.1 Basic principles

N-dimensional matching consists of putting into registration series of n-dimensional “images” (n typically being 3 or 2). The general idea of matching consists first of extracting some key features in one of the (n-dimensional) images, for example using segmentation-like methods. Correspondence is then established by searching for this restricted number of features in the other images (56). The major categorization of matching methods is between rigid and nonrigid (elastic) matching, and between registration using specific landmark pairs (points, lines, specific patches) and fully volumetric techniques (e.g. using hierarchical correlation (57)).

The key geometrical features used as landmark pairs must be invariant under translations

and rotations of the objects and sometimes under elastic deformations (58). In other words, they should ideally be intrinsic object characteristics. The features that can be used are:

- pixels (or voxels): they are numerous and therefore difficult to match between slices. However, their intensity is usually invariant when seen from different angles;
- line segments: characterized by more attributes than pixels (e.g. length, orientation), henceforth easier to match (matching based on 3D space curves is described in (59),e.g.). However, not all contours are invariant;
- regions: less numerous and even easier to match, but more difficult to extract and usually not invariant;
- critical points, such as vertices, cusps, saddles, etc.: difficult to extract, but intrinsic to the objects and consequently very robust features for matching.

6.2 Registration of 3D images

When images come from several devices, *multimodality matching* by *registration techniques* allows the user to have simultaneous access to various types of information; examples of applications are *functional* or *metabolic imaging*, where functional information is mapped onto 3D anatomical structure (60). A related type of matching occurs when the digital information is to be compared with more generic knowledge, such as provided by an anatomical atlas (56).

In robotics, the problem of stereo pairs registration for extracting depth information from 2D images is rather well solved; solutions have also been proposed for matching 3D geometric objects (61). This is not the case in 3D medical imaging, where in addition to the complexity brought by the third dimension, the input data does not have the nice geometrical regularities offered by man-made objects. Various approaches are being investigated, either for rigid or elastic matching. *Rigid matching* is of course much simpler, since objects shape is fixed; the only degrees of freedom concern the respective displacement of the objects to be matched (62,63).

Elastic matching is much more realistic (and more complex), since it takes into account possible shape variations. One of the current research trends consist of extracting the desired invariant features by means of local quadric approximations of the object surface (64). This rather local approach is to be somewhat opposed to more global techniques where the objects to be matched are iteratively deformed; the optimal alignment is the one minimizing an ad-hoc cost function (58,65).

6.3 Analysis of 2D image sequences

When an *image sequence* is provided, i.e. a series of images (usually 2D) from the same device at different times, *temporal matching* provides correspondence between features from the different images. If the time interval between successive acquisitions is large, the

matching process is conceptually similar to multimodality matching: a predefined structure is tracked along time (e.g. (17,66)).

However, when this interval is small, other techniques can be employed since it can be assumed that the modifications between images are minute; in such case, the features to be matched have a size of the order of a pixel. The recovery of the *optical flow*, which is the estimate of the 2D projection of the 3D velocity field of the scene, is an essential step. After its calculation, it is possible to segment objects based on their intrinsic motion: different objects will have different optical flow vectors. This segmentation can be used to measure the temporal evolution of objects characteristics as well as to extract some structural information (67,68,69).

The main approaches to the computation of the optical flow are the brightness constancy method, which assumes that the brightness of a given point is constant between two frames (70), the spatial gradient constancy method, stating that the spatial luminance gradient is stationary over time (69), the local correlation approach, which looks for correspondence between groups of nearby points (68), and the energetic approach, which uses 3D directional filters in {2D space, time} that have maximal response along motion directions (71). It can be noted that the first three methods correspond respectively to tracking pixels, edges or regions between successive images. It is nowadays possible to compute estimates of the optical flow in almost real-time. However, despite being used more and more in computer vision applications, these techniques have probably not shown enough maturity to attract medical imaging researchers.

7. Recognition

Object recognition involves two phases, namely learning and exploitation. During the *learning phase*, typical characteristics of the objects to be recognized are determined and stored (such as morphological measures). This determination is accomplished by techniques such as described in this paper, i.e segmentation and features extraction. Learning is often supervised: the user identifies which are the relevant, discriminating characteristic features.

During *exploitation*, the unknown objects are passed through the same pipeline as the one used for learning; consequently, the object characteristics are measured. These measures are then compared with the learned ones; the result indicates which is the object considered.

The most classical recognition paradigm is *statistical* (26): covariance matrices are computed for the discriminative measures, usually under the multivariate Gaussian assumption. Parametric discriminant functions are then determined, allowing classification of unknown samples. These functions lead to the use of distances in parameter space (Euclidean, Maha-

nalobis, quadratic): an unknown sample, i.e. a point in parameter space, is classified with the closest class center according to the selected distance. Many recognition programs are based on this approach (e.g. (72,73)), whose major problems are the need for large training samples, the difficulty of taking structural information into account, and the issue of defining an appropriate parameter space.

Neural networks have been advocated as excellent classifiers, due on the one hand to their ability to define non-linear decision surfaces, on the other hand to the back-propagation learning mechanism which alleviates the need for explicitly defining the parameters space (28,29). Many applications of the neural paradigm to medical imaging are appearing, not only for classification but also for segmentation, viewed as a classification to be made between object and background (e.g. (74)). The major problems with this approach are also the need for large training sets and the impossibility of taking structural information into account. Despite these drawbacks, the neural approach is very much in favour for its rather unsupervised nature (although there is still the need for a training set).

Structural and symbolic approaches are the most commonly encountered ones (1,3), although purely symbolic approaches (27) are not as common as a decade ago. The basic idea is to describe the objects by means of symbolic structures, and to compare the observed structure to a model one; these structures can have varying levels of complexity, as enumerated in §6.1. Typical of symbolic approaches are comparison by graph matching, or by using grammars; in the latter case, an object is described by a string of symbols, according to a given grammar, which is then parsed by a lexical analyser. This paradigm allows to take into account structural information and does not require large training samples. Its major drawbacks are the difficulty to process data in order to extract symbols, and consequently a difficulty to cope with noise. Actually, the problem with symbolic approaches exemplifies the major difficulty of computer vision: how to extract symbolic information from numeric data.

8. Trends

8.1 Medical image analysis and Computer vision

Many of the techniques described in this paper are now routinely used in biomedical laboratories. Relying on these rather classical methods, various trends are emerging. New detectors are being developed, which should lead to less invasive and more precise imaging modalities. Functional and multi-modal imaging is becoming commonly used, for discovering and understanding functional structures. It is possible to further combine data in order to incorporate medical knowledge, either from a patient file or from a medical atlas.

A lot of effort is put into the realization of integrated hospital information systems (HIS), and more particularly related to medical imaging, into PACS development (14). Although at first glance one might feel that the issue is essentially technological, there still are many theoretical problems to solve: image coding for storage and transmission (f.ex. using wavelets), software development, design of common standards for storing and displaying data, artificial intelligence techniques for “intelligent” images data bases creation.

Feature-based indexing techniques are being developed, that will be essential for retrieval in image data bases. Such data bases should allow information access by semantic content, and therefore make possible queries such as “find all radiographs showing a broken arm”. A related challenge is *knowledge acquisition*. Although various methods exist for supervised learning, as exposed above, unsupervised extraction of the key characteristics that will enable subsequent recognition is a problem far from being solved.

New surgical techniques are characterised by the trend towards minimally invasive therapy using new and more precise tools (e.g. endoscopic interventions, laser surgery). A patient will benefit from this development by less risk, reduced pain, shorter hospitalization and faster recovering. These new techniques can be decisively supported by a detailed preoperative planning and by precise feedback information based on image analysis.

Further, sophisticated robotic vision methods will permit the advent of *medical robotic*s, that is the design of robots able to perform complex surgical interventions in an automated manner (21).

All these developments require more sophisticated computer vision techniques. Amongst the future developments will certainly come an increase in the use of *top-down information*, at least to circumvent the limits reached by the current segmentation methods (e.g. (20,75)). An initial segmentation of the 3D image will be performed; the result will then be compared with symbolic medical knowledge either from an atlas or from the patient’s file. This will allow, through a feedback loop, to better the initial anatomical segmentation and finally label the organs.

Other progresses stemming from computer vision will certainly be used in medical imaging. One of the major problem in scene analysis is computational complexity: there is an infinity of possible mappings from object models to the digitized scene. Various approaches have as a primary objective to decrease the size of the solution space, for example by using aspect graphs, active vision, or perception-based approaches such as focus of attention (5,76). There is also a trend towards reduction of the amount of data acquired by using non-regular grids, such as polar lattices that mimic a human retina by providing high-resolution only near the center.

8.2 The Medical Imaging Workstation of the (Near) Future

The physician’s workstation of the (near) future will have three main characteristics: it will

be reasonably “intelligent”, extensively connected and highly interactive.

Reasonable intelligence will be provided by computer vision and “artificial intelligence” techniques. However, progress is not as fast as practitioners would hope. Despite some attempts at providing guidance using expert systems (77), it seems highly unlikely that user’s experience will soon be totally replaced by artificial means. As a consequence, interaction (and “real intelligence” (78)) will remain necessary. In a first stage, machine intelligence will mostly be used to assist users in image segmentation and feature extraction. Rule based algorithms for image interpretation can further help in efficient image analysis.

Connectivity, both physical (networking) and in spirit, will be required for integration into a Hospital Information System. It will also allow to distribute more evenly (financial) resources: a low cost workstation will be used for the interaction with the physician while at the same time the CPU intensive image analysis with all computations and interaction control will be done on another, more powerful computer. It will also put to practice the recent workplace concepts revolving around groupware environments, which aim at verifying the Gestalt law “the sum is more than the parts”.

Interaction will allow the user to perform sophisticated processing while still retaining a control over the results, and also to inspect in three-dimensions the results of the analysis. The interaction will less and less be performed using keyboard and mouse, more and more using *virtual reality* devices. Depth images will be recreated by means of stereoscopic imaging, using shutter glasses or helmets with miniaturized video screens. It will be possible to mix images with sounds, using multimedia techniques (see other articles in this special issue). All this will allow physicians to interactively explore the body (and even the soul!) of their patient.

8.3 The European COVIRA project

This section discusses a large project which incorporates some of the main premises recognized as important issues for future research in medical image analysis, namely interdisciplinary collaboration, clinical tests and validation. As a further requirement, standardization of software tools has to ensure interchangeability and transportability.

A European Community research project within the AIM (Advanced Informatics in Medicine) program is titled COVIRA (COmputer VIsion in RAdiology) (79). The objective of COVIRA is to realize a large software system providing efficient computer assistance in neuroradiological diagnosis, stereotactic and open neurosurgical planning and radiation therapy planning. The project has started in January 1992 and will run over 3 years.

COVIRA is a cooperative effort of 3 leading industrial partners (IBM, Philips, Siemens), 8 academic research groups and 6 clinical institutions in Belgium, Germany, Italy, the Netherlands, Spain, U.K., Crete and Switzerland (ETH-Zürich, G.G.).

The software system will be based on rigorously specified standards incorporating object oriented programming. All algorithms will be implemented in a portable manner following the upcoming ISO standard on Image Processing and Interchange (IPI) and will enter a common pool of computer algorithms for medical multimodality 2-D and 3-D image analysis.

The main functional capabilities of the system will be:

- 2-D and 3-D image visualization;
- 2-D and 3-D image segmentation combining automatic segmentation algorithms with interactive image editing capabilities;
- reconstruction of cerebral vascular system based on the combination of 3D MR Angiography (MRA) and 2-D Digital Subtraction Angiography data (DSA);
- multimodality image registration;
- digital and annotated anatomical atlas of the human head.

The unifying objective of the project is to provide and clinically test pilot application systems for multimodality image analysis for diagnosis and therapy management, based on commercially available workstation and accelerator hardware with standard software tools. The system will be tested and validated by the clinical partners of the COVIRA consortium.

The collaboration between academic, industrial and clinical partners provides a unique combination of strengths and combines expertise in the fields of image analysis research, software engineering, manufacturing of medical equipment and, last but not least, diagnosis and therapy. It is the hope of the consortium that the project will finally prove clinical usefulness and cost effectiveness of computer assistance in various application fields.

9. Summary and conclusion

Is there any motivation for incorporating computer vision algorithms into a multimedia workstation for physicians? The primary function of a workstation in the medical milieu is to provide otherwise inaccessible information, in order to help the physician interpret, predict and plan. The motivation is therefore strong to incorporate computer vision algorithms into such environment, since they provide tools and methods for extracting the necessary pieces of knowledge. By showing the range of computer vision methods in medicine, the purpose of this paper was to convince the reader that the answer to the above question should be positive.

Parallel to the recent progress in computer vision and feature extraction, there is also an increased complexity of the images generated in medicine. More and more imaging modalities tend to generate very large sets of images containing tremendous amounts of complex

information. Physicians can hardly master all the complexity of this information and it is necessary to assist them with appropriate tools for (multimodal) image interpretation.

There is a final motivation for burdening the overloaded physician with the requirement of mastering yet another set of methods and tools. It is easy to be impressed by more attractive and meaningful pictures, without paying too much attention to the data manipulation involved and the consequences on the resulting images. This is why we feel there is a need for a better grasp of the potential provided by image analysis and computer vision methods.

10. Acknowledgments

This work is supported in part (for T.P.) by grants from the Swiss National Fund for Scientific Research (FNRS 20-33467.92, 4023-27036).

11. References

[1] Rosenfeld, A.; Kak, A.C. Digital Picture Processing; 2nd Edition. Academic Press; 1982.

[2] Foley, J.D.; Van Dam, A.; Feiner, S.F.; Hughes, J.F. Computer Graphics: Principles and Practice; 2nd Edition. Addison-Wesley; 1990.

[3] Ballard, D.H.; Brown, C. Computer Vision. Englewood Cliffs, NJ: Prentice Hall. 1982.

[4] Marr, D. Vision. San Francisco: Freeman. 1982.

[5] Wechsler, H. Computational Vision. Academic Press; 1990.

[6] Three-Dimensional Imaging. Vannier, M.W.; Marsh, J.L.; Biondetti, P.R., Ed. Special Issue of: Computerized Medical Imaging and Graphics. 12(1); Jan.-Feb. 1988.

[7] Visualisation of 3D medical images. Special Issue of: IEEE Computer Graphics and Applications. March 1990.

[8] Magnenat-Thalmann, N.; Thalmann, D. New Trends in Computer Animation and Visualization. J. Wiley. 1991.

[9] 3D Imaging in Medicine. Fuchs, H.; Hohne, K; Pizer, S. Eds. NATO ASI Series, Springer-Verlag; 1990.

[10] Dhawan, A.P. Review: biomedical image processing and future trends. Computer. Methods and Programs in Biomedicine. 31(3-4):141-183. 1990.

[11] Prewitt, J.M.S. Parametric and nonparametric recognition by computer: an application to leukocyte image processing. Advances in Computers, 12:285-414; 1972.

[12] Kunt, M., Ikonomopoulos, A., Kocher, M. Second generation image coding techniques. Proc. IEEE. 73(4):549-574; Apr. 1985.

[13] Picture Archiving and Communication Systems. Huang, H.K., Ed. Special Issue of: Computerized Medical Imaging and Graphics. 15 (3); May-Jun. 1991.

[14] Medical Communications. McGarty, P.T.; Blaine, G.J.; Goldberg, M., Eds. Special Issue of: IEEE J. on Selected Areas in Communications. 10(7); Sep. 1992.

[15] Herman, G.T. Image Reconstruction from Projections: the Fundamentals of Computed Tomography. Academic Press. 1980.

[16] Kak, A.C.; Slaney, M. Principles of Computerized Tomographic Imaging. IEEE Press; 1988.

[17] Gerig, G.; Kikinis, R.; Kuoni, W.; von Schulthess, G.K.; Kübler, O. Semiautomated ROI analysis in dynamic MRI-studies, Part I: Image analysis tools for automatic correction of organ displacements. J. of Comp. Assisted Tomography. 15(5):725-732. 1991.

[18] Ayache, N.; Boissonat, J.D.; Cohen, L.; Geiger, B.; Levy-Vehel, J.; Monga, O; Sander, P. Steps towards the automatic interpretation of 3D images. In: (9): 107-120.

[19] Székely, G.; Brechbühler, Ch.; Kübler, O.; Ogniewicz, R.; Budinger, T. Mapping the Human Cerebral Cortex using 3-D Medial Manifolds. In: Robb, R.A., Ed. Visualization in Biomedical Computing VBC'92. Proc. SPIE. 1808:130-144. Oct. 1992.

[20] Dhawan, A.P.; Juvvadi, S. Knowledge-based analysis and understanding of medical images. Computer. Methods and Programs in Biomedicine. 33(4):221-239. 1990.

[21] Benabid, A.L.; Lavallée, S.; Hoffmann, D.; Cinquin, P.; Demongeot, J.; Danel, F. Potential use of robots in endoscopic neurosurgery. Acta Neurochirurgica. Suppl. 54:93-97. 1992.

[22] Suetens, P.; Fua, P.; Hanson, A.J. Computational strategies for object recognition. ACM Computing Surveys. 24(1):5-61. Mar. 1992.

[23] Pavlidis, T. Algorithms for Graphics and Image Processing. Computer Science Press; 1982.

[24] Pun, T.; Blake, E. Relationships between image synthesis and analysis: towards unification? Computer Graphics Forum. 9(2):149-163; Jul. 1990.

[25] Serra, J. Mathematical Morphology and Image Analysis. Academic Press; 1988.

[26] Duda, R.O.; Hart, P.E. Pattern Classification and Scene Analysis. John Wiley and sons; 1973.

[27] Fu, K.S. Syntactic Pattern Recognition and Applications. Prentice Hall. 1982.

[28] Rumelhart, D.E.; McClelland, J.L.; and the PDP Research group. Parallel Distributed Processing. Vol.1: Foundations. 1986.

[29] Poggio, T.; Girosi, F. Networks for approximation and learning. Proc. IEEE. 78(9):1481-1497. Sep. 1990.

[30] Frank, J.; Verschoor, A.; Boublik, M. Computer averaging of electron micrographs of 40S ribosomal subunits. Science. 214:1353-1355. 18 Dec. 1981.

[31] Illingworth, J.; Kittler, J. A survey of the Hough transform. Computer Vision, Graphics and Image Processing. 44:87-116. 1988.

[32] Haralick, R.M.; Sternberg, S.R.; Zhuang, X. Image analysis using mathematical morphology. IEEE Transactions on Pattern Analysis and Machine Intelligence. 9(4):532-550. Jul. 1987.

[33] Besl, P.J. Geometric modeling and computer vision. Proc. IEEE. 76(8):936-958.

Aug. 1988.

[34] Pentland, A.P. Perceptual organization and the representation of natural form. Artificial Intelligence. 28:293-331. May 1986.

[35] Kass, M.; Witkin, A.; Terzopoulos, D. Snakes: active contour models. In: Proc. First Int. Conf. on Comp. Vision, London, 259-268. Jun. 1987.

[36] Cohen, I.; Cohen, L.D.; Ayache, N. Using deformable surfaces to segment 3D images and infer differential structures. Computer Vision, Graphics and Image Processing - Image Understanding. 56(2):242-263. Sept. 1992.

[37] Lutz, R.; Pun, T.; Pellegrini, C. Colour displays and look-up tables: real time modification of digital images. Computerized Medical Imaging and Graphics. 15(2):73-84. 1991.

[38] Deriche, R. Fast algorithms for low-level vision. IEEE Trans. Pattern Analysis and Mach. Intelligence. 12(1):78-87. Jan. 1990.

[39] ter Haar Romeny, B.M.; Florack, L.M.; Koenderink, J.J.; Viergever, M.A.; Scale space: Its natural operators and differential invariants. Information Processing in Medical Imaging, Colchester, A.C.F.;Hawkes, D.J. Eds. Lecture Notes in Computer Science, Springer-Verlag. 511:239-255. 1991.

[40] Fleck, M.M. Some Defects in Finite-Difference Edge Finders. IEEE Trans. Pattern Analysis and Mach. Intelligence. 14(3):337-345. March 1992.

[41] Ohanian, P.P.; Dubes, R.C. Performance evaluation for four classes of tectural features. Pattern Recognition. 25:819-833. 1992.

[42] Jain, A.K.; Farrokhnia, F. Unsupervised texture segmentation using Gabor filters. Pattern Recognition. 24(12):1167-1186. 1991.

[43] Mallat, S. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans. Pattern Analysis and Mach. Intelligence. 11(7):674-693. Jul. 1989.

[44] Rasure, J.R.; Williams, C.C. An integrated visual language and software development environment (Khoros). J. of Visual Languages and Computing. 2:217-246. 1991.

[45] Jacot-Descombes, A.; Todorov, K.; Pellegrini, C.; Hochstrasser, D.F.; Pun, T. LaboImage: a workstation environment for research in image processing and analysis. Computer Applications in the Biosciences. 7(2):225-232. Apr. 1991.†

[46] Ligier, Y.; Funk, M.; Ratib, O.; Perrier, R.; Girard, C. The OSIRIS User Interface for Manipulating Medical Images. Proc of NATO ASI meeting on “Picture Archiving and Communication System (PACS) in Medicine. Evian, France, Oct 1990.

[47] Trus, B.L.; Unser, M.; Pun, T.; Steven, A. Digital image processing of electron micrographs: the PIC System II. Scanning Microscopy Supplement 6, in press.

[48] Sander, P.; Zucker, S.W. Inferring surface trace and differential structure from 3D images. IEEE Trans. Pattern Analysis and Machine Intelligence. 12(9). Sept. 1990.

[49] Monga, O.; Deriche, R.; Rocchisani, J.-M. 3D edge detection using recursive filtering: application to scanner images. Computer Vision, Graphics and Image Processing. 53(1):76-87. Jan. 1991.

[50] Boissonnat, J.-D. Shape reconstruction from planar cross-sections. Comp. Vision,

†. This program is available by anonymous ftp; please contact the first author for further information.

Graphics and Image Processing. 44:1-29. 1988.

[51] Stytz, M.R.; Frieder, G.; Frieder, O. Three-dimensional medical imaging: algorithms and computer systems. ACM Comp. Surveys. 23(4):421-499. Dec. 1991.

[52] Lorensen, W.; Cline, H. Marching cubes: a high resolution 3D surface reconstruction algorithm. Computer Graphics. 21:163-169. 1987.

[53] Hoehne, K.R.; Bernstein, R. Shading 3D images from CT using grey-level gradients. IEEE Trans. Medical Imaging. 5(1):45-47. 1986.

[54] Robb, R.A. A software system for interactive and quantitative analysis of biomedical images, In: (9):333-361.

[55] Harris, L.D. Display of multidimensional biomedical image formation. In: Three-Dimensional Biomedical Imaging. Robb, R.A., Ed. 2nd Vol. Boca-Raton, FL: CRC Press. 1985. 125.

[56] Bookstein, F.L. Principal warps: Thin-plate splines and the decomposition of deformations. IEEE Trans. PAMI. 11:567-585. 1989.

[57] Collins, D.L.; Peters, T.M.; Dai, W.; Evans, A.C. Model based segmentation of individual brain structures from MRI data. In: Robb, R.A., Ed. Visualization in Biomedical Computing VBC'92. Proc. SPIE. 1808:10-23. Oct. 1992.

[58] Bajcsy, R.; Lieberson, R.; Reivich, M. A computerized system for elastic matching of deformed radiographic images to idealized atlas images. J. of Computer Assisted Tomography. &(4):618-625. Aug. 1983.

[59] Gueziec, A.; Ayache, N. Smoothing and matching of 3-D space curves. In: Robb, R.A., Ed. Visualization in Biomedical Computing VBC'92. Proc. SPIE. 1808:259-273. Oct. 1992.

[60] Schäfer, M.; Scheppelmann, D.; Meinzer, H.P. Multimodal segmentation of medical images. In: Proc. IAPR - Workshop on Machine Vision Applications. Tokyo. Dec. 7-9, 1992.

[61] Lowe, D.G. Fitting parametrized three-dimensional models to images. IEEE Trans. Pattern Analysis and Machine Intell. 13(5):441-450. May 1991.

[62] Arun, K.S.; Huang, T.S.; Blostein, S.D. Least squares fitting of two 3D sets. IEEE Trans. Pattern Analysis and Mach. Intell. 9(5):698-699. Sep. 1987.

[63] Pellizzari, C.A. et al. Accurate 3D registration of CT, PET and MR images of the brain. J. of Comp. Assisted Tomography. 13:20-26. 1989.

[64] Ayache, N.; Boissonat, J.-D.; Brunet, E.; Cohen, L.; Chièze, J.P.; Geiger, B.; Monga, O.; Roccisani, J.M.; Sander, P. Building highly structured volume representation in 3D medical images. Computer Aided Radiology. 89:765-772. 1989.

[65] Schad, L.R.; Boesecke, R.; Schlegel, W.; Hartmann, G.H.; Sturm, V.; Strauss, L.G.; Lorenz, W.J. Three-dimensional image correlation of CT, MR, and PET studies in radiotherapy treatment planning of brain tumors. J. of Computer Assisted Tomography. 11(6):948-954. Nov./Dec 1986.

[66] Herlin, I.; Ayache, N. Features extraction and analysis methods for sequences of ultrasound images. In: Goos, G.; Hartmanis, J., Eds. Proc. of the Second Europ. Conf. on Comp. Vision, ECCV'92, Santa Margherita Ligure, Italy, May 1992. Springer-Verlag, Lecture Notes in Computer Science. 588. 1992.

[67] Horn, B.; Schunk, B.G. Determining optical flow. Artificial Intelligence. 17:185-203. 1981.

[68] Zhang, Z.; Faugeras, O.D. Determining motion from 3D line segments: a comparative study. Image and Vision Computing. 9(1):10-19. Feb. 1991.

[69] Verri, A.; Straforini, M.; Torre, V. Computational aspects of motion perception in natural and artificial vision systems. Proc. Phil. Trans. In press. 1992.

[70] Hildreth, E.C. The computation of the velocity field. Proc. R. Soc. London B. 221:189-220. 1984.

[71] Heeger, D.J. Optical flow from spatiotemporal filters. In: Proc. 1st Int. Conf. on Computer Vision. New York: IEEE. 181-190. 1987.

[72] Gerig, G.; Martin, J.; Kikinis, R.; Kübler, O.; Shenton, M.; Jolesz, F.A. Automatic segmentation of dual echo MR head data. In: Proc. 12th Int. Conf. Inf. Proc. in Medical Imaging IPMI'91. Colchester, A.C.F.; Hawkes, D.J., Eds. Lecture Notes in Computer Science 511. 175-185. Springer Verlag. 1991.

[73] Pun, T.; Hochstrasser, D.F.; Appel, R.D.; Funk, M.; Villars-Augsburger, V.; Pellegrini, C. Computerized classification of two-dimensional gel electrophoretograms by correspondence analysis and ascendant hierarchical clustering. Applied and Theoretical Electrophoresis. 1(1):3-9. 1988.

[74] Kippenhan, J.S.; Barker, W.W.; Pascal, S.; Nagel, J.; Duara, R. Evaluation of a neural network classifier for PET scans of normal and Alzheimer's disease subjects. The J. of Nuclear Medicine. 33(8):1459-1467. Aug. 1992.

[75] Kita, Y. Model-driven contour extraction for physically deformed objects - Application to analysis of stomach X-ray images. Proc. 11th IAPR. The Hague, The Netherlands. Vol. 1:280-284. Aug. 30 - Sep. 3, 1992.

[76] Milanese, R.; Bost, J.-M.; Pun, T. A bottom-up attention system for active vision. In: Proc. 10th. Europ. Conf. on Artificial Intell. ECAI 92, Vienna. J. Wiley. August 3-7, 1992.

[77] Hu, Z.; Pun, T.; Pellegrini, C. An expert system for guiding image segmentation. Computerized Medical Imaging and Graphics. 14(1):13-24. 1990.

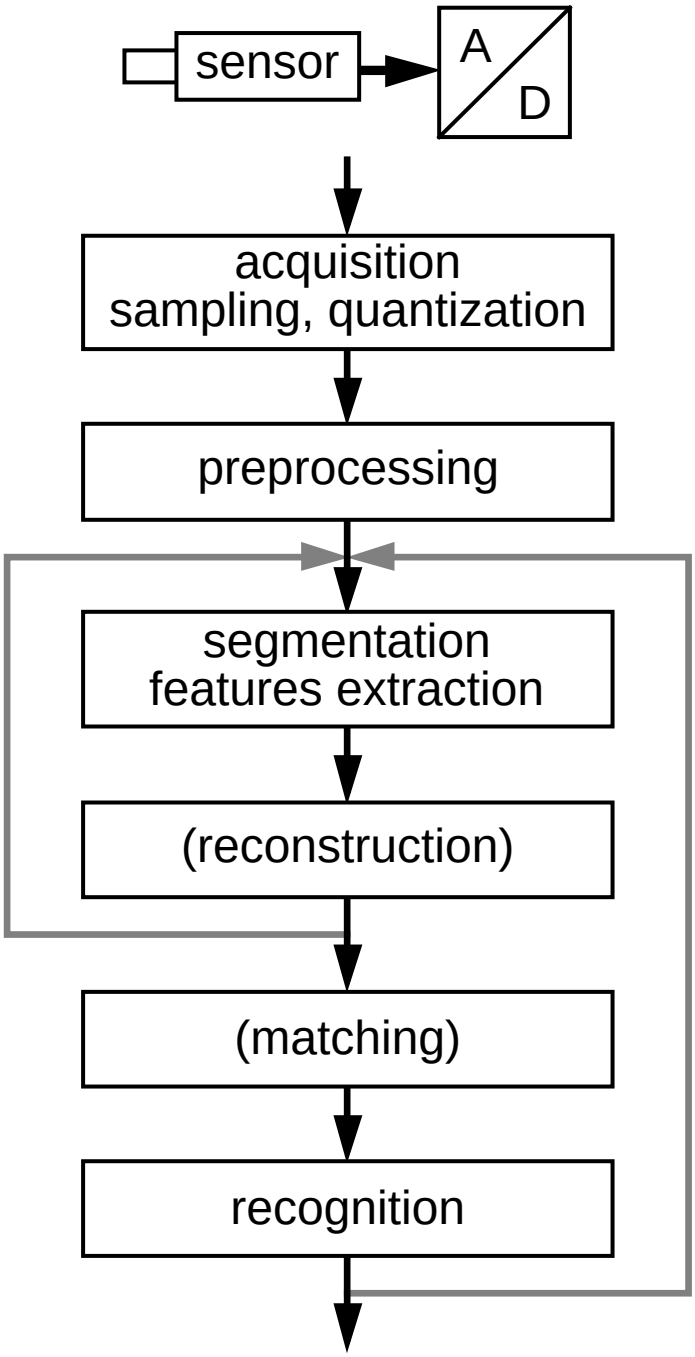
[78] Ledley, R.S.; Ayers, W.R. Editorial. Computerized Medical Imaging and Graphics. 12(1):v-xviii. Jan.-Feb. 1988.

[79] COVIRA, Computer Vision in Radiology. European Community project within the Advanced Informatics in Medicine (AIM) framework, number 11659. Project duration: Jan. 1992 - Dec. 1994.

Figure captions.

Figure 1: The generic medical image analysis pipeline.

Figure 1


sensor
A
D
acquisition
sampling, quantization
preprocessing
segmentation
features extraction
(reconstruction)
(matching)
recognition