The ESPRIT LTR Research Project: "Nonlinear Model-Based Analysis and Description of Images for Multimedia Applications (NOBLESSE)"

V. Pahor*, G. Ramponi*, R. Castagno[†]

* D.E.E.I - University of Trieste
via A. Valerio,10 -I-34127 Trieste, Italy
e–mail: pahor@ipl.univ.trieste.it, ramponi@univ.trieste.it

[†] Signal Processing Laboratory, Swiss Federal Institute of Technology
CH–1015 Lausanne, Switzerland
e–mail: castagno@lts.de.epfl.ch

1 Introduction

The project "Nonlinear Model-Based Analysis and Description of Images for Multimedia Applications, Noblesse" is an ESPRIT Long Term Research action funded by the Commission of the European Communities. It has started in January 1996 and has a duration of three years. In this project, nonlinear techniques are developed that enable a model-based symbolic description of images and image sequences, to support tasks in multimedia applications. Novel results are expected in several fields of multimedia application such as image archiving, image sequence coding and video sequence processing.

The members of the Noblesse consortium are: Prof. H. Burkhardt (Coordinator), Albert Ludwigs University of Freiburg, Germany; Prof. M. Gabbouj, Tampere University of Technology, Finland; Prof. M. Kunt, Swiss Fed. Inst. of Technology Lausanne, Switzerland, Dr. S. Marshall, University of Strathclyde Glasgow, United Kingdom; Prof. I. Pitas, Aristotle University of Thessaloniki, Greece; Prof. G. Ramponi, University of Trieste, Italy; Prof. J. Serra, Prof. F. Meyer, Ecole des Mines, Fontainebleau, France.

2 Objective and goals

The objective of the project is to develop new nonlinear model-based descriptions of images and image sequences. Indeed, the large amount of image-data in multimedia applications makes it necessary to code the information in a model-based symbolic form, and a high measure of compactness in the data representation can be foreseen if the potential of nonlinear models is taken into account. Beside the applications for image coding purposes, high-level description is an absolute prerequisite to support the application of higher level functions like model-based browsing and navigation, keying, image sequence interpolation, tracking and

finding salient regions, querying (similarity measures, indexing, fuzzy similarity measures). In order to reach these goals, long-term research is carried out in areas like optimal model adaption, segmentation, region identification, image and scene interpretation, symbolic description, motion estimation etc. Current models were mainly developed for image coding purposes. They are rather simple and far away from being optimal and do not contribute to more complex tasks like those needed in image databases.

The research is focusing on standard images and video sequences; however, more advanced tasks for future applications such as stereo pairs and 3D image processing and virtual reality are investigated.

3 Scientific content

To meet the challenges of the project tasks, the partners share their expertise in image processing, nonlinear filtering, pattern recognition, image coding etc., focusing on the potential of *nonlinear models*, such as: polynomial models, motion models, set-based models (rank-order, morphological, region growing like watershed etc.), stochastic and chaotic models, genetic algorithms and fuzzy models, canonical frames with invariants, 3D models, models for interpolation.

Polynomial models. Polynomial models have proved to be efficient for several tasks, such as object description, evaluation of invariants, determination of classification boundaries, data interpolation, data processing, etc., which are essential in multimedia applications. As a consequence, polynomial models are often used as a first-level preprocessing technique. A potential application is the preprocessing of document images with polynomial filters for enhancement purposes. Another interesting application is in the field of the enhancement (e.g. edge-preserving noise smoothing, blocking effect removal) in still and moving images. Finally, we can consider edge extraction as a basic low-level technique which is preliminary to a wide class of segmentation algorithms.

Model–based analysis of 3D models. One of the key elements of future multimedia systems is the acquisition, analysis, description, recognition and rendering of 3D objects. The Noblesse project addresses the investigation of novel promising techniques in 3D object analysis and recognition based on nonlinear techniques coming from the morphology, neural networks and invariance theory.

Motion models. The analysis and the modelling of motion in video sequences represents a key issue in image processing. Second generation video coding techniques, based on the identification and the manipulation of homogeneous regions and semantically meaningful objects, are currently under development. This further increases the need for accurate modelling of the motion characteristics of the scene. Nonlinear techniques have proved to be efficient in this field thanks to their flexibility and their adaptation to the nonlinear characteristics of the problem.

Fuzzy models. Fuzzy techniques have a very good chance to become an effective resource for multimedia data processing since fuzziness really affects many aspects of multimedia environments: multisensory data can be noisy and incomplete and decision making often has to face the fusion of the information deriving from different sources. In this project, modelling based on fuzzy rules is addressed to image data analysis and processing as well. Specific application areas include segmentation, enhancement, filtering and data fusion.

Optimisation techniques and self-learning structures. Optimisation techniques and self-learning structures represent the backbone of any processing applied to images and image sequences. A large number of ad-hoc structures is available in the literature, but the lack of a solid analytic basis often limits their use to very specific applications. In parameterised processing models in which exhaustive search is prohibitive, optimisation techniques become essential in finding those best "operators" according to a given performance criterion. When these algorithms are made adaptive, in the classical filtering theory sense, or self-learning, they become even more suitable for multimedia environments in which scenes are inherently nonstationary.

Models for pyramidal and interpolative processing. Pyramidal data structures show promising performances, while having some interesting properties, in problems related e.g. to data fusion or multimedia applications. The choice of the filtering operator before decimation or after interpolation is not trivial, especially when nonlinear filters such as morphological and order statistics operators are considered.

Interpolation is needed to generate missing images in sequences (coding, clock change, time sampling); interpolation is also required for archiving documents, and restoring them at various scales (still or moving images); finally, interpolation is involved in the production of cartoons, as well as in generating zooms, or virtual reality.

The conversion of TV archive material to high definition, larger scale formats presents problems of blurring and colour bleeding. Non linear filters such as stack and morphological filters may be developed and used to sharpen the image and correct colour errors.

Object recognition, region classification. In order to support tasks like searching and browsing in image databases it is necessary to synthesise a symbolic image description. A first step towards this goal is the identification of objects in a given scene.

Techniques for region classification are important e.g. for image coding applications. Conventional transform coding relies on a static partition of the image into blocks. This results in blocking artifacts in the decoded image. Region based coding tries to circumvent this problem by modelling an image as a combination of non-stationary visual primitives (e.g. object contours and textures) which are tied to the objects in the scene.

Set-based models, like rank-order, morphological, region growing like watersheds etc. Segmentation of images and image sequences is a low-level processing step which is very crucial in any further analysis of the underlying information. The consortium is applying and analysing the recently developed watershed and other region-growing based methods to segment multimedia images and image sequences. A major task is the extension of the existing algorithms to segment 3-D data.

Non-linear techniques, developed by several teams in this consortium among others, have proved to be appropriate for restoration of images and superior to linear techniques. These include Boolean and stack filter related operations. An interesting approach is to model images by using Hidden Markov Chains for the purpose of filter optimisation for restoration application.

4 Noblesse activity

4.1 Technical content and first results achieved

The scientific work planned within Noblesse is organised in seven work packages (abbreviated by WP). In the following, for each of them the tasks planned and the specific technical commitments are described. Moreover, the most significant results achieved during the first year are summarised.

WP 1 – Object recognition, region classification. The major focus of this WP is the development of methods to support tasks in image databases and coding like data retrieval, intelligent search and browsing, navigation and keying. Some key issues are invariants for object recognition, object recognition by using canonical frames, object recognition from 3D data sets, recognition and tracking of objects in image sequences, region classification, texture classification by using invariant gray scale features, identification of regions in image sequences, image segmentation in colour TV images.

During the first year, a general methodology for the construction of invariant gray scale features has been devised which is applicable to arbitrary transformation groups. Specialised algorithms have been developed for the group of image rotations and translations, the affine group and the group of projective image transformations. These groups are of primary interest for computer vision applications to describe viewpoint and pose dependency. The invariance properties, the robustness and the separation properties of the invariant gray scale features have been experimentally investigated.

Moreover, a novel, analytical tool for the description, unique affine normalization and pose determination of segmented, general, volumetric 3D objects through the formation of complete moment invariants has been developed.

The problem of "colour bleeding" is one application of the recognition techniques being developed.

WP 2 – Optimisation techniques and self-learning structures. Novel adaptive and self-learning techniques for processing mainly images in a multi-media environment were proposed to be investigated in this WP: we mention in particular stochastic algorithms in design of nonlinear filters, optimal adaptive nonlinear filtering algorithms, unified stochastic optimisation algorithms, genetic algorithms for optimal morphological related filters, genetic algorithms for optimal rule-based fuzzy systems.

Optimal design procedures have been developed for Boolean, stack, and FIR–Boolean hybrid filters for realizing the prediction stage in lossless grey-level image compression. A training framework has been proposed to design optimal nonlinear filters for various signals (images of a natural scene) and image processing tasks (e.g. restoration). Boolean based filters are shown to be robust, making it possible to build libraries of "optimal" filters, which are suitable for a set of applications.

Two novel techniques, a vector quantisation based locally adaptive structure for stack filtering and a Library–Median–Stack ($\mathcal{L}$–M–S) filter have been introduced and tested revealing good performances in image restoration applications.

Several optimisation schemes, such as simulated annealing and genetic algorithms, have been investigated for the search of soft morphological filters having an optimal performance in a given signal processing task.

A method which had been previously developed for the optimisation of standard (structural) morphological filters has been applied to various real world tasks including restoration of corrupted film and video material as supplied by the BBC. In this method a genetic algorithm is used to optimise the choice of the basic morphological operations.

WP 3 – Motion models. During the period of interest, the development of motion models drew a considerable amount of interest in the video processing community, especially in relation to the recent development in the field of object-based video coding.

These issues are addressed in WP3 of Noblesse with special regard to motion based segmentation of image sequences, object tracking, interpolation and model–based stereo image coding. To achieve this goal, several techniques are under investigation in order to improve the robustness of motion based modelling, such as multi–feature image analysis, Bayesian criteria, Radial Basis functions, mathematical morphology.

In particular, a tool for unsupervised spatio–temporal video sequence segmentation and object tracking has been developed and used in a second generation object based coding scheme. Radial Basis functions have been applied to the problem of segmentation of motion fields.

Based on the Bayesian criteria, hierarchical scale–space algorithms have been applied in order to estimate the depth field, the optical flow, the surface gradient and the motion parameters from gray–scale images. Eventually, a connected operator has been proposed that allows to detect objects, whose movement is fast enough for their position to be disconnected along the time axis.

WP 4 – Data fusion. This work package is concerned with the fusion of image and speech data. Some key issues are the development of a database incorporating speech information and image data, the utilisation of the database to synchronise lip movements with speech signals, the integration of transmission of image and speech data, the development of robust tracking algorithms based on speech and image data for real time location of facial features and the development of new image segmentation techniques for location of mouth area.

The principal contribution in the first year has been to develop and maintain a speech and image relational database to be used in the area of data fusion. The database contains a sequence of mouth images and the corresponding speech data both in raw and coefficient form. The objective of this work is to establish and quantify the relationship between the speech and image data for improved recognition, compression and synchronisation of both data streams.

WP 5 – Model-based analysis of 3D objects. Three types of techniques for analysing 3D objects are investigated: *morphological techniques for 3D image segmentation*, based on 3D watershed-types of algorithms or on 3D morphological Voronoi tesselations; *morphological techniques in 3D object modelling*, extended from 2D schemes (based either on skeleton or on morphological decomposition) for the description of 3D shapes; *Neural techniques for 3D image segmentation*, where Radial Basis Functions (RBF) are used to model 3D objects with proper choice of the basis functions.
So far, several algorithms have been developed:
- a novel efficient algorithm for incorporating mathematical morphology into the Iterative Closest Point (ICP) algorithm (used to solve 3-D registration problems).
- algorithms that compute morphological Voronoi tessellation as well as watershed transforms on large 3D images, image regions and image surfaces that cannot be stored on RAM. They have very small memory requirements and are reasonably fast in contrast to classical algorithms that fail in this case.
- a novel dynamic link architecture that employs morphological shape decomposition as feature extraction mechanism, used to analyse a facial image region and successively to verify the identity of each person from a training set.

The 3D binary shape decomposition has also been studied: the computational complexity has been analysed and some implementation and optimisation aspects have been discussed.

WP 6 – Polynomial models, fuzzy models. The key issues of this WP are polynomial models for image enhancement and segmentation, polynomial splines for image processing, polynomial-based invariant image features, fuzzy models for image analysis and processing, fuzzy methods for nonlinear estimation and filtering.

The so called *Rational Filters* have been introduced. First, a nonlinear operator which is able to effectively attenuate the noise which corrupts an image, while introducing small distortions on the image details, has been described.

Based on a similar concept, a novel interpolation scheme which effectively tackles the problem of accurate image interpolation has been developed. Finally, the Rational Filter approach has been used for removing blocking artifacts in coded image sequences. Polynomial models have also been used to develop nonlinear unsharp masking schemes.

The research work on fuzzy techniques has mainly focused on the development of new FIRE (Fuzzy Inference Ruled by Else-action) operators, dedicated to signal and image processing applications. A new class of these operators, based on adaptive piecewise linear fuzzy sets, has been applied to the impulse noise removal. A method for the automatic generation of the fuzzy rulebase, based on Genetic Algorithms, has been also developed.

Research has been also done on the connection of Volterra theory with algorithms for the construction of invariant features, on B–Splines for modelling noisy multivariate data and on fuzzy estimators for image processing (e.g. the Fuzzy Cell Hough Transform has been introduced as contour detector).

WP 7 – Set-based models. Four major application areas are studied: old movies restoration, confocal microscopy, segmentation of flame images and conversion from usual TV format to HDTV.

In year one, mathematical morphology has been applied to the restoration of old movies. Motion picture defects have been classified on the basis of the method that will be used for their restoration (global, local and local-non-moving defects). Preliminary work on restoration of global defects has been performed (i.e. histogram correction along the sequence and reduction of image jittering).

For the detection of local defects, two methods have been proposed and tested. This work was generalised and applied to the detection of fast moving objects, occlusions and new objects in image sequences. Moreover, in order to interpolate the data that was destroyed by the local defects, morphological tools (spatial dilations and temporal closings) have been employed as well as the extrapolation of Fourier coefficients calculated around the defect inside the missing region.

4.2   Interaction between Work Packages

The combination of the techniques developed in the separate work packages is essential. In order to strengthen the cohesion of the consortium and to stimulate the cooperation between the partners, the following applications are considered: region based image coding (WP1,WP3, WP6, WP7), synchronisation between the lip movements and the speech signals in video coding (WP4, WP6, WP7), old movies restoration (WP2, WP4, WP7), searching in image databases (WP1,WP5, WP6).

5   Potential industrial applications

The Noblesse project is dedicated to the development of new model-based methods to support tasks in multimedia applications.

A major application area for the results of the Noblesse project is videophone technology. For example, lip synchronisation techniques became important due to the increase in compression rates and improved picture quality. Applications can be foreseen also in the film industry for the dubbing of foreign language movies, in the synthesis of computer graphics and humans in film and animation. It may also have applications in aides for the deaf.

The description of video sequences in terms of joint spatio–temporal models has a wide range of applications. In addition to segmentation and coding applications (addressed in particular by MPEG-4), compact video representation is a key issue in applications involving also indexing and annotation. Many applications involving for example indexing and querying of images and videos will require in the future indexing capabilities based on the static and dynamic pictorial contents.

Simple and effective algorithms for the enhancement and for the segmentation of images are a very important resource in the growing field of low bit-rate image coding. The results of the Noblesse project in this area might be of interest to enterprises which are developing systems for teleconferencing, security and telesurveillance. Similarly, the deliverables produced in the area of document image enhancement might be of interest for a wide range of enterprises involved in the design of systems ranging from PC-based optical character recognition software packages to graphics vectorisation systems for engineering drawings and maps, to automatic mail sorting systems. The final users include financial and banking companies, hospitals, administrative offices and the postal service.

All partners in Noblesse consortium have links to several industrial enterprises (Nokia, Hewlett Packard, Thomson, Alcatel, Philips, British Telecom and others). These links represent possible channels for transferring the results of the Noblesse project to applications.

Finally, we have to mention the particular interest expressed by several companies for a specific application considered in the Noblesse framework, that is old motion picture restoration.

6   Noblesse publications

In the first year, the Noblesse project has produced a consistent output, characterised by 11 journal papers, 36 conference papers, 14 technical reports, 2 PhD theses on Noblesse related topics and 1 World Wide Web image/speech database. The list of publications, the related abstracts and also some full papers in Postscript version are available on the main Noblesse WWW home page http://www.informatik.uni-freiburg.de/~noblesse/. Other interesting information is available on the partners' WWW pages easily reachable from the main Noblesse home page.