MUSCLE

Network of Excellence

Multimedia Understanding through Semantics, Computation and Learning

Project no. FP6-507752

Bi-monthly progress report # 21: July- August 2007

Due date of deliverable: 15.10.2007
Actual submission date: 16.10.2007

Start date of project: 1 March 2004   Duration: 48 months

Name of responsible editor(s): Patricia Ho-Hune (ERCIM)

version 1.0

	Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)	
	Dissemination Level	
PU	Public	✓
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Contents

1 Overview activities in WP1

1.1 General scientific and administrative coordination

Activities & Achievements

- Administrative and financial coordination of the network
- Organisation of regular audio-conferences.
- E-Teams reorganization and showcase extension monitoring
- Participation to CIVR 07 on 9-11 July in Amsterdam where MUSCLE showcases were demonstrated
- Modifications of MUSCLE financial cost statements and Audit certificates for period 3 requested by the Commission's service.
- Version 3 of MUSCLE Annex I including JPA4 was sent to the project officer
- Maintenance of MUSCLE website by ERCIM
- Reimbursement of MUSCLE integration expenses (fellowship, sponsorship)
- Preparation of MUSCLE participation to the IBC exhibition in Amsterdam in September (www.ibc.org).
- Preparation of MUSCLE participation to the CeBIT Eurasia exhibition in Amsterdam in October (www.cebitbilisim.com).
- Cooperation with related European projects

2 Overview activities in WP2

2.1 Contribution by CEA LIST

Researchers involved

Pierre-Alain MOELLIC

Activities & Achievements

Organization of the ImagEVAL Workshop, 12 July 2007, University of Amsterdam. CEA LIST, as the scientific animator of the evaluation campaign organized a workshop, one day after the CIVR Conference. During this workshop, the participants have described their works and systems used for the ImagEVAL evaluation challenges. Discussions have been planned for the future of the campaign. The presentations and papers presented during the workshop have been plublished on the ImagEVAL website.

Events

ImagEVAL Workshop, University of Amsterdam

Publications

See http://www.imageval.org/e_publications.html

2.2 Contribution by ISTI-CNR

Researchers involved

Patrizia Asirelli, Massimo Martinelli, Ovidio Salvetti, Marco Tampucci

Activities & Achievements

The work focused on further developing the 4M Infrastructure. An investigation activity for assessing different policies for uploading XML and RDF documents that detail the features extracted from images or their regions and their annotations has been completed. According to the result of the investigation activity the policies for uploading XML and RDF documents have been modified. Now the features extracted from an image and their regions are stored into the same XML document (into the proper collection) and the annotations are stored into an RDF document which has the same univocal name of XML document. In the frame of ET1 E-Team work, a refinement activity regarded the user interface, which is, indeed, under constant improvement. In particular, new operations for viewing extracted MPEG-7 features from image regions and for deleting XML documents have been added.

Events

Participation to W3C MMSEM XG

2.3 Contribution by UPC

Researchers involved

Ferran Marques, Xavier Giro

Activities & Achievements

As a contribution to the e-team "Choosing Features for CBIR" UPC has developed a first version of a manual annotation tool that uses a hierarchical region-based representation of an image based on a previously computed Binary Partition Tree. The tool reads an MPEG-7 XML file that contains a description of the Binary Partition Tree as well as the paths to the location of the actual image and its labelled partition. Once this is done, the application reads from disk both image and partition and parses the MPEG-7 XML file to initialize the process. As a result, the application shows the image on screen and activates a region-based navigation system through the BPT. The user can then select any pixel from the image as an anchor point and the region associated to the corresponding BPT leaf is selected. Upper nodes in the hierarchical representation can be selected or unselected using the mouse wheel. The selection of regions is visualized by making them transparent on an overlay mask of selectable colour and transparency. The tool allows the selection of multiple regions to be annotated as an instance of the semantic class selected in the menu on the left. Semantic classes can be simple

(one part) or composite (multiple parts). In the second case, the annotation is performed by selecting on the image the regions associated to each of its parts. Semantic classes shown on the menu can be loaded from disk or interactively defined using the interface itself. As a result of the whole annotation process, the tool generates an MPEG-7 compliant XML file containing those regions from the initial partition that correspond to instances of the considered semantic classes.

2.4 Contribution by TU VIENNA-PRIP

Researchers involved

Allan Hanbury, Lech Szumilas

Activities & Achievements

Work on the preparation of two tracks at the ImageCLEF 2007 campaign has continued. These are ImageCLEFphoto, the photographic image retrieval task, and the object retrieval task. During July and August, the runs sent in by the participants were evaluated and papers giving overviews of the tasks and results were written.

Preparation of the MUSCLE / ImageCLEF workshop on Image and Video Retrieval Evaluation to be held on the 18th of September 2007 in Budapest has continued. The deadline for submitting papers was the end of July. 10 papers were submitted, of which 9 were accepted for presentation.

The MUSCLE live image retrieval event took place at the CIVR in Amsterdam. 3 groups participated (all are MUSCLE members). The event was judged to be a success by the participants. A paper describing the event was written to be presented at the MUSCLE / ImageCLEF workshop.

Events

The MUSCLE live image retrieval event took place at the CIVR in Amsterdam on Tuesday, July 10.

Publications

A paper describing the live image retrieval event was written to be presented at the MUSCLE / ImageCLEF workshop.

Papers describing the ImageCLEFphoto and ImageCLEF object retrieval tasks and results were written to be presented at the ImageCLEF workshop on the 19th-21st of September.

2.5 Contribution by IBaI

Researchers involved

Petra Perner Horst Perner

Activities & Achievements

We have been working on implementing the interface between the prototype-based classification and all the other tools in the Muscle E-Team. Testing the interface and the methods based on multimedia data bases. Reimplementing method and evaluating again. Working on a test and evaluation report. We have been working on implementing the interface between the prototype-based classification and all the other tools in the Muscle E-Team. Testing the interface and the methods based on multimedia data bases. Reimplementing method and evaluating again. Working on a test and evaluation report.

2.6 Contribution by UTIA

Researchers involved

Haindl, M.

Activities & Achievements

Information of image quality and / or performance of image restoration methods can be expressed by evaluative criteria. The classical criteria are used for rather unrealistic image evaluation while the original (undegraded) image is available. More difficult case is in image restoration assessment when the original is unknown. Blind criteria serve as an image quality estimator based on the restored image. Criteria are tested and compared on the synthetic degraded data set (the "ideal" is known) to set out the most appropriate criteria for the real astronomical images.

Publications

MP-588

3 Overview activities in WP3

3.1 Contribution by Trinity College Dublin

Researchers involved

Rozenn Dahyot, Gerard Lacey

Activities & Achievements

We propose a method to restore colonoscopy videos that have low quality RGB images. The main problem concerns a time delay occurring in between the recordings of the R, G and B colour channels. As the camera is moving along in the colon, sometimes quickly, the resulting images show non properly matched R, G and B causing blurry effects that impede the medical doctors or computer-aided analysis methods. We proposed to restore this artefact by first

equalizing the colour channels and then performing a robust camera motion estimation and compensation. Experimental results show significant improvements from the original videos.

Publications

Restoration of colour channel misalignments in colonoscopy videos Rozenn Dahyot and Gerard Lacey, Computer Science technical report TCD-CS-2007-27, Trinity College Dublin, 2007.

3.2 Contribution by UvA

Researchers involved

Jasper Uijlings, Nicu Sebe, Cees Snoek

Activities & Achievements

- Organization of ACM International Conference on Image and Video Retrieval, July 2007 - Organization of the VideOlympics event as part of the Muscle Evaluation Showcase - Color Interest Points for Image Retrieval In image retrieval scenarios, many methods use interest point detection at an early stage to find regions in which descriptors are calculated. Finding salient locations in image data is crucial for these tasks. Observing that most current methods use only the luminance information of the images, we investigate the use of colour information in interest point detection. A way to use multi-channel information in the Harris corner detector is explored and different colour spaces are evaluated. To determine the characteristic scale of an interest point, we presented a new colour scale selection method. We show that using colour information and boosting salient colours results in improved performance in retrieval tasks. An article will be published in ICIP 2007.

Events

ACM International Conference on Image and Video Retrieval, July 2007

3.3 Contribution by UNIS

Researchers involved

B Goswami, a Khan, WJ Christmas

Activities & Achievements

Lip categorization experiments: we are using SIFT features to build a generic object detector type approach to lip segmentation, a Viola-Jones style Haar feature based lip detection system We performed a comparative analysis of a Minimum Covariance Estimator based lip segmentation system with a system that used Level Sets and SVM approaches to pixel-by-pixel lip segmentation. This is in preparation for collaboration with Universitat Politècnica de Catalunya as part of the body analysis e-team.

Events

Bud Goswami attended the summer school on Biometrics in Alghero, Sardinia

3.4 Contribution by UCL

Researchers involved

Fred Stentiford, Rob Shilson, Shijie Zhang

Activities & Achievements

Work Package 3.2 – Visual Saliency
Attention based motion algorithms tend to attach importance to the boundaries of moving objects and ignore the non-salient homogeneous regions within objects. The current algorithms have been extended by introducing a growing algorithm that assigns motion vectors to object interiors providing there is a colour match with the boundary regions and that the resulting displacement obtains a match with the succeeding frames. Papers have been presented at the London Communications Symposium 2007 and the International Conference on Image Processing 2007.

Work on models of Human Visual Accommodation has been presented at the London Communications Symposium 2007. Initial results presented show that a visual attention model provides a better model for the subjective assessment of out-of-focus images. Little research has been done in this field to date and more comprehensive experiments are planned to support a journal publication.

Publications

S Zhang and F Stentiford, "Region Growing for Motion Segmentation using an Attention Based Algorithm," London Communications Symposium, 2007.

S Zhang and F Stentiford, "Motion Detection using a Model of Visual Attention," International Conference on Image Processing, San Antonio, September 2007.

R Shilston and F Stentiford, "Preliminary Subjective Focus Assessment Results," London Communications Symposium, 2007.

3.5 Contribution by ISTI-CNR

Researchers involved

Umberto Barcaro, Sara Colantonio, Davide Moroni, Ovidio Salvetti, Anna Tonazzini

Activities & Achievements

A method for the automatic computation of Left Ventricle (LV) Ejection Fraction (EF) from echocardiographic images sequences has been developed. EF is a fundamental parameter for

heart function assessment, routinely estimated by sonographers during transthoracic echocardiography examinations. However, being based on manual contour tracing, the computation of EF is time consuming and prone to intra and inter observer variability. For these reasons, we addressed the problem of automatic segmentation of the LV cavity in ultrasound image sequences. Indeed after having segmented the LV cavity, estimation of the EF boils down to the application of the conventional Simpson's rule. The segmentation procedure exploits a variational formulation of Level Set methods, i.e. methods in which the evolving contour is represented implicitly as the zero level set of a suitable 2D scalar function. With the aim of solving the difficult initialization problem, we augmented the Level Set method by a computational stage that provides a suitable initialization contour: in fact, level set methods, as other active contour schemes based on energy minimization, may become trapped in undesirable local minima. The initial contour is obtained by means of the introduction of mimetic criteria, based on the idea of mimicking the processing steps applied by an expert observer for the identification of LV cavity. Besides, we employ a recent variational formulation of level sets without contour re-initialization, thus circumventing the subtle problems encountered in the application of the classical level set initialization PDE. In the framework of e-team 4 on "Shape modeling", we developed some experimental software in the MATLAB environment and we tested it on several echocardiographic image sequences. The method has been reported in a paper orally presented at the Open German Russian Workshop 2007. Future research (again related to e-team 4 on "Shape modeling") will aim at obtaining more accurate volume computations by combining different echocardiographic views. Indeed this issue is crucial for the evaluation of "misshapen" ventricles, for which the geometric assumptions underlying Simpson's rule are no longer valid. Further, our activity on statistical analysis of microspectroscopy signals for algae classification and phylogenetic comparison has been presented at the conference Mass-Data Analysis of Images and Signals in Medicine, Biotechnology and Chemistry - MDA 2007 (Leipzig, July 2007).

Events

-1- Participation to the Open German Russian Workshop 2007 on Pattern Recognition & Image Understanding (Ettlingen, Germany, 20-23 August 2007)

2- Participation to the conference Mass-Data Analysis of Images and Signals in Medicine, Biotechnology and Chemistry - MDA 2007 (Leipzig, July 2007)

Publications

Barcaro U., Moroni D., Salvetti O. - Left ventricle segmentation in ultrasound sequences for the automatic computation of ejection fraction. In: Open German Russian Workshop on Pattern Recognition & Image Understanding (Ettlingen, Germany, 20-23 August 2007), Forshungsinstitut fur Optronik und Mustererkennung, 2007

3.6 Contribution by MTA SZTAKI

Researchers involved

S. Fazekas, D. Chetverikov (SZTAKI), T. Amiaz, N. Kiryati (TAU-Visual)

Activities & Achievements

A joint (SZTAKI-TAU) paper on dynamic texture detection and segmentation was prepared by S. Fazekas, D. Chetverikov, T. Amiaz and N. Kiryati, for submission to a top computer vision journal. The paper will summarise the MUSCLE research on the topic. For this submission, additional experiments were carried out that compare the original level-set based method with the graph-cut based variant as well as a number of simplified fast variants of the algorithm. For the fast variants, a multiresolution technique for optical flow calculation was implemented and tested that allows for robust setting of a critical parameter of the detection algorithm.

Problems

Possible bugs in the public-domain graph cut software, leading to some uncertainty in the experimental results. Solution: postpone publication of detailed description, give preliminary results only.

3.7 Contribution by TUG

Researchers involved

Martina Uray, Helmut Grabner, Peter Roth, Horst Bischof

Activities & Achievements

1. We started doing some work concerning robust LDA learning. Due to the fact that our constructed subspace not only contains discriminative information but some additional reconstructive information it is possible to apply the main ideas of robust PCA leaning. For proof of concept the first experiments are based on the assumption that occlusions are known, such that only undisturbed pixels are considered for creating the subspace.

2. We improved our person detection framework making it faster and more stable.

3.8 Contribution by UPC

Researchers involved

Camilo Dorea Chang, Ferran Marques, Montse Pardas

Activities & Achievements

The PhD thesis "Hierarchical Partition-Based Representations of Motion-Coherent Regions For Video Object Segmentation" was defended at UPC. As member of the jury, Dr. Patrick Bouthemy, from IRISA/INRIA VISTA, assisted to the presentation. This thesis addresses the development of hierarchical partition-based representations for the extraction of video objects. Motion, as well as spatial features, from multiple frames are used to generate a set of regions structured within a hierarchy of scale and motion coherency. The resulting representations offer a global description of the entire video sequence and enhance semantic analysis potential.

3.9 Contribution by Advanced Computer Vision GmbH - ACV

Researchers involved

Csaba Beleznai

Activities & Achievements

We developed a new computationally efficient clustering approach which is capable to delineate objects in multistructured two-dimensional distributions by closed convex polygonal contours. The clustering step obeys to a stability criterion locating maximally stable covariance matrix estimates over scales, followed by a spatial grouping scheme generating convex hulls. The validity of the method has been confirmed in the applied context of a pedestrian counting system. The clustering method is scale-adaptive and it can be used to segment multiple foreground objects from background in presence of noise and clutter in a stable manner.

Events

C. Beleznai gave a tutorial presentation at the SSIP'07 (Summer School on Image Processing) in Szeged, Hungary.

3.10 Contribution by AUTH

Researchers involved

C. Cotsaces, N. Nikolaidis, I. Pitas

Activities & Achievements

AUTH continued the work on query-by-example video retrieval using semantic signatures incorporating face-related information video. The major advantage of such signatures is that they are highly invariant to nearly all types of distortion. The proposed method (developed earlier in the project) uses the pre-extracted output of face detection and recognition to characterize the video through pulse series that describe the appearance intervals of persons in the video (e.g. actors) and perform fast semantic query-by-example retrieval. Apart from the experiments on a synthetic database that have been conducted in the past, experimental evaluation of the method on a database of real video data (movies and TV series) using the results of a commercial face detection/tracking and recognition software has been performed. The results showed that the method can work efficiently in real situations.

Publications

C. Cotsaces, N. Nikolaidis, I. Pitas, "Face-based digital signatures for video retrieval and fingerprinting", IEEE Transactions on Circuits and Systems for Video Technology, accepted for publication.

3.11 Contribution by GET-ENST

Researchers involved

Beatrice Pesquet-Popescu, Christophe Tillier

Activities & Achievements

We prepared the presentation for the accepted paper at ICIP2007 on the three-step non-linear lifting scheme. In parallel, work continued on the perturbation analysis of this scheme. Another research direction was the investigation of a clustering by compression method, which is based on finding similar parts in a lossless compressed representation of signals. A trainee started to work on this topic, with application to image indexing, and possibly to video.

3.12 Contribution by TU VIENNA-PRIP

Researchers involved

Allan Hanbury, Lech Szumilas

Activities & Achievements

Work on object recognition using image keypoints based on a measure of symmetry combined with a new feature describing the shape of the area around keypoints has continued. These features have also been applied to encoding salient point features in medical X-ray images in combination with the optimisation algorithms described in the previous bi-monthly report (see also the two publications listed below). Lech Szumilas has intensified cooperation with the University of Amsterdam by spending July doing research there (in total 3 months).

Publications

Donner R., Micusik B., Langs G., Szumilas L., Peloschek P., Friedrich K. and Bischof H.: Object Localization Based on Markov Random Fields and Symmetry Interest Points, International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Brisbane, Australia, 2007

Donner R., Micusik B., Langs G. and Bischof H.: Sparse MRF Appearance Models for Fast Anatomical Structure Localisation, British Machine Vision Conference (BMVC), Warwick, UK, 2007

3.13 Contribution by IRIT-UPS

Researchers involved

Christine Sénac, Elie El Khoury, Philippe Joly

Activities & Achievements

The segmentation method (GLR/BIC combination) previously designed for audio segmentation in the SAMoVA team and presented during ICASSP 2007 is now used with adapted parameters in the context of visual data. The first results obtained with 24 hours of French TV broadcasting are promising. The segmentation is hierarchical, without a priori knowledge and without backtracking: a first step permits to detect the shots boundaries then a second step (using the same segmentation method) permits to detect the broadcasts boundaries. Experiments have to be pursued to validate this method of macro segmentation of the TV stream.

3.14 Contribution by INRIA Ariana

Researchers involved

Josiane Zerubia, Ian Jermyn, Avik Bhattacharya, Peter Horvath, Ting Peng, Aymen El Ghoul.

Activities & Achievements

Ting Peng, joint PhD student of INRIA Ariana and the LIAMA Institute in Beijing, has continued to work on models for road network segmentation from very high resolution (0.5m) satellite images. As described in the previous report, Ms Peng has developed three models for the extraction of narrower roads. Two involve asymmetrizing the interaction in the higher-order energy term to allow it to vary with the relative orientation of the contour normal vectors. The third uses a hitherto unrecognized linear term that takes into account the relative orientation of the vector between the two interacting points and the normal vectors. The idea of all three models is to enforce road edge straightness over a larger range than the road width, although they do this in different ways. After obtaining good results with the first, nonlinear model, Ms Peng moved on to experimental testing of the new linear model, since this is less intensive computationally and is therefore preferable if the quality of the results is comparable. The results obtained are again very good, comparable in quality to the nonlinear model, and it seems likely, subject to confirmation in further tests, that this will become the model of choice for further development. At this point, one model suffices to extract both wide and narrow roads, by varying the model parameters. The next step is to use a single model to extract roads of different widths simultaneously. There are two clear possibilities for this: one is to change the interaction function defining the model to allow a broader ranger of widths; the other is to use a multiscale approach to extract large roads first, gradually moving to narrower roads as the scale is decreased. Aymen El Ghoul, student at Sup'Com, Tunis, is nearing the end of his second internship with INRIA Ariana as a prelude to taking up a PhD position in the group in September. He is working on extending the higher-order active contour model of road networks, and its phase field formulation, to the cases of grid-like road networks and river networks, with the aim of extracting these entities from high resolution satellite and aerial images. As described in the previous report, as part of his internship work, he performed a stability analysis for a long straight bar analogous to that performed by Mr. Horvath for a circle, and tested it against the results of gradient descent with success. Subsequently, he has performed the same type of stability analysis for the new nonlinear model developed by Ting Peng. This analysis is important for Mr El Ghoul's work because the new model can be used either to encourage straighter network segments and perpendicular junctions, suitable for grid networks, or to encourage more undulating network segments and looser junctions, suitable for river networks. He is currently working on an INRIA Research Report that will summarize his internship work, and serve as a basis for his doctoral research.

Peter Horvath, joint PhD student of INRIA Ariana and the University of Szeged, submitted his thesis to the University of Szeged in August, and will submit it to the University of Nice Sophia Antipolis in September. He will defend in December. Avik Bhattacharya, joint PhD student of INRIA Ariana and ENST, is writing up his thesis, and will submit it to the University of Nice Sophia Antipolis in October. All being well, he will defend in December.

3.15 Contribution by INRIA-Vista

Researchers involved

Alexandre Hervieux, Patrick Bouthemy

Activities & Achievements

We continue the work on original statistical trajectory-based approaches addressing several issues related to dynamic video content understanding: unsupervised clustering of events, recognition of events and detection of unexpected events. Appropriate local differential features combining curvature and motion magnitude are defined and robustly computed on the motion trajectories. These features are invariant to image translation, in-the-plane rotation and the spatial scaling. The temporal causality of the features is then captured by Hidden Markov Models whose states are properly quantized values. The similarity between trajectories is expressed by exploiting the HMM framework. We have evaluated our methods on several data sets including typical classes of synthetic trajectories (with noise), and real trajectories obtained from sports videos. We are also processing video surveillance data sets. We have favorably compared our method to other ones, including feature histogram comparison, the use of the longest common subsequence (LCSS) distance and SVM-based classification. A part of this ongoing work has been recently presented at ICIP'07.

Publications

"A HMM-BASED METHOD FOR RECOGNIZING DYNAMIC VIDEO CONTENTS FROM TRAJECTORIES" (2007), Alexandre Hervieu, Patrick Bouthemy, Jean-Pierre Le Cadre; in Proc. ICIP'07, Texas, USA.

3.16 Contribution by INRIA-Vista

Researchers involved

Ivan Laptev

Activities & Achievements

Within the MUSCLE showcase project "Content-Based Copy Detection for Videos and Still Images" we designed an efficient implementation of space-time interest point detection. The new implementation enables real time video processing and the computation of video descriptors for on-line demonstration of video copy detection techniques. In the collaboration with INRIA-Imedia, space-time interest points (STIP) have been evaluated as an alternative

video description in the context of video copy detection. The comparison of STIP with alternative methods performed by INRIA-Imedia has been recently presented at CIVR'07.

Publications

"Video Copy Detection: a Comparative Study" (2007), J. Law-To, L. Chen, A. Joly, I. Laptev, O. Buisson, V. Gouet-Brunet, N. Boujemaa, and F.I. Stentiford; in Proc. CIVR'07, Amsterdam, The Netherlands, pp. 371-378.

3.17 Contribution by ENSEA

Researchers involved

Prof. Sylvie Philipp-Foliguet, Philippe-Henri Gosselin, Julien Gony

Activities & Achievements

We participated to the Technovision ImagEval live contest during CIVR 2007, July 9-12, Amsterdam, where we got the best recall price. Two kinds of queries were asked: image category retrieval from keywords, and keyword retrieval from images. Each challenger had 10 minutes per query to provide its results. In order to deal with these queries, we enhanced our retrieval system to handle text and image retrieval. We use kernel function-based techniques to merge the different features, which allow us to use our active learning methods for interactive learning. This live contest showed the interest of our interactive techniques where good results had to be returned in a short time for a large multi-media database.

3.18 Contribution by UTIA

Researchers involved

Haindl,M. and Vacha,P.

Activities & Achievements

We developed fast and robust image retrieval measures for CBIR systems that utilise novel illumination invariant features extracted from three different Markov random field (MRF) based texture representations. These measures allow retrieving images with similar scenes comprising colour textured objects viewed with different illumination brightness or spectrum. The proposed illumination insensitive measures are compared favourably with Local Binary Patterns, steerable pyramid and Gabor textural features and demonstrated in the illumination invariant recognition of textures from the Outex database. Our CBIR demonstration was presented in the MUSCLE & 6th ACM International Conference on Image and Video Retrieval, Amsterdam.

Publications

MP-591, MP-592

3.19 Contribution by UTIA

Researchers involved

Haindl,M. Grim,J. Mikes,S.

Activities & Achievements

A multispectral texture defect detection method based on the underlying three-dimensional spatial probabilistic image model was developed. The model first adaptively learns its parameters on the flawless texture part and subsequently checks for texture defects using the recursive prediction analysis. We provide colour textile defect detection results that indicate the advantages of the method.

Publications

MP-587

4 Overview activities in WP4

4.1 Contribution by Trinity College Dublin

Researchers involved

Rozenn Dahyot, Conor Kelly, Gavin Kearney

Activities & Achievements

The use of multiple audio streams from digital mixing consoles is presented for application to real-time enhancement of synchronised visual effects in live music performances. The audio streams are processed simultaneously and their temporal and spectral characteristics can be used to control the intensity, duration and colour of the lights. The efficiency of the approach is tested on rock and jazz pieces. The result of the analysis is illustrated by a visual OpenGL 3-D animation illustrating the synchronous audio-visual events occurring in the musical piece.

Publications

Visual enhancement using multiple audio streams in live music performance, Rozenn Dahyot, Conor Kelly, Gavin Kearney, 31st Audio Engineering Society International conference, London, June 2007

4.2 Contribution by TAU SPEECH

Researchers involved

PIs: David Burshtein, Arie Yeredor

Activities & Achievements

1) Support vector machine training for improved hidden Markov modeling: The algorithm was extended to the case of connected speech. We used an N-best list to produce an initial set of hypothesized word strings and then rescored them using an SVM composite re-scoring model. We also designed a training method for the SVM re-scoring model in this scheme and started conducting experiments on actual connected speech data. This work is also related to WP6.

2) Single channel audio separation: We examined several Linear Programming approaches, such as Simplex and interior point methods in order to find the most suitable method for implementation of the single-channel separation learning stage (might be made more efficient if a few "intelligibility-related constraints" are incorporated, such as continuity of the coefficients). Matlab implementation of the algorithm was programmed using Matlab's linear programming solver. A dynamic method was incorporated for finding an efficient step-size for the gradient-descent (as large changes in the dictionary matrix can have a devastating effect on the signs of elements of the coefficient vector, which are all supposed to be non-negative).

4.3 Contribution by CEA LIST

Researchers involved

Gregory Grefenstette

Activities & Achievements

CEA LIST continued research on natural language models for image recognition. We have been experimenting on calculating prior co-occurrence probabilities for objects found in images, from pure text sources, and from human annotated images. Our first results, showing that these two independent sources have the same distribution (and thus that you can use text statistics for image processing) was submitted to the International Symposium on Visual Computing and accepted for presentation there in Lake Tahoe, Nevada in November 2007. The final version of the paper has been uploaded to the MUSCLE website: http://www.muscle-noe.org/images/DocumentPDF/2007-09-14%20ISVC%20Pitel-Millet-Grefenstette%20Comparing%20Prior%20Text%20Image-finalv2.pdf

Publications

Deriving A Priori Co-occurrence Probability Estimates for Object Recognition from Social Networks and Text Processing, Guillaume Pitel, Christophe Millet, Gregory Grefenstette, ISVC, Lake Tahoe, Nov. 2007

4.4 Contribution by AUTH

Researchers involved

E. Benetos, C. Kotropoulos

Activities & Achievements

Automatic musical instrument identification using a variety of classifiers is addressed. Experiments have been performed on a large set of recordings that stem from 20 instrument classes. Several features from general audio data classification applications as well as MPEG-7 descriptors are measured for 1000 recordings. The first classifier is based on non-negative matrix factorization (NMF) techniques, where training is performed for each audio class individually. Several NMF variants are utilized besides the standard NMF method, such as the local NMF and the sparse NMF. In addition, 3-layered multilayer perceptrons, normalized Gaussian radial basis function networks, and support vector machines employing a polynomial kernel have also been tested as classifiers.

Events

C. Kotropoulos attended 2007 4th Sound & Music Computing Conf. held at Lefkada in July 2007.

4.5 Contribution by CNR-ISTI

Researchers involved

Graziano Bertini, Massimo Magrini, Leonello Tarabella

Activities & Achievements

- Preparation of a presentation-demo for eNTERFACE '07 Summer Workshop (Istanbul 2007) regarding results on our researches in interactive computer music based on HW/SW tools developed at Signal and Image Lab of CNR-ISTI: in particular two original gesture recognition devices and systems, or hyper-instruments, have been used (PalmDriver and Handel) together with the "pCM" real-time music language based on C-language, for sound synthesis and events management. The presentation has been accompanied with a L. Tarabella live performance using these tools and systems (title: "Gesture touchless live computer music" by computerART project of S&I Lab ISTI-CNR). - Going on testing and up-date of ARIA Dynamic Sound (release2) algorithm for transients and dynamic range enhancement of compressed music. - Collaboration to the preparation of MODEM EU project final reporting.

Events

Partecipation to eNTERFACE '07 the SIMILAR NoE Summer Workshop on Multimodal Interfaces (July 16th - August 10th, 2007), Bogazici University, Istanbul, Turkey.

Publications

Two chapters (in two languages) of the MODEM EU project Final Book: 1) description of the didactic activities of Signal and Images Lab od ISTI in the field of musical signals synthesis and processing. 2) Description of the MODEM web-site online musical sequencer, developed by ISTI.

4.6 Contribution by IRIT-UPS

Researchers involved

Hélène Lachambre, Julien Pinquier, Régine André-Obrecht

Activities & Achievements

In a context of audio indexing and retrieval, we have worked on improving a speech activity detector. This new detector is based on the previous detector developed in our team, and based on 4 Hz modulation energy and modulation entropy. A threshold was applied on these parameters to detect the presence of speech. In our new detector, we apply two thresholds on both parameters, to have two « certainty » zones, and one « uncertainty » zone (between the thresholds). Our work has been to find new ways to deal with the « uncertainty » zone: learn a GMM on both certainty zones (of speech and non speech) to estimate presence or not of speech in the uncertainty zone.

4.7 Contribution by TU Vienna - IFS

Researchers involved

Thomas Lidy, Jakob Frank, Andreas Rauber

Activities & Achievements

We integrated our Audio Analysis and Feature Extraction System with a Polyphonic Transcription System and a Symbolic Feature Extractor in a joint effort with University of Alicante (Spain), to build an integrated system for improved Music Genre Classification. After extensive tests on reference music databases we prepared the system for participation in the 2007 Music Information Retrieval Evaluation eXchange (MIREX) and submitted it there. Research on Audio Segmentation was finalized: We developed a system that is able to find segment boundaries and to cluster audio segments from the same type. Work on the web service enabling training and creation of Self-Organising (Music) Maps continued.

Events

MIREX (Music Information Retrieval Evaluation eXchange) 2007, Vienna, Austria, September 23 -27, 2007

5 Overview activities in WP5

5.1 Contribution by INRIA - Texmex

Researchers involved

Patrick Gros, Gwenole Lecorve, Stephane Huet, Pascale Sebillot, and Guillaume Gravier

Activities & Achievements

The WP5 has decided to edit a book about multimedia processing techniques and interfaces. INRIA texmex contributed to two chapters of this book. The first of them, written by Manolis Delakis, Guillaume Gravier and Patrick Gros is devoted to stochastic models for video structuring. It summarizes the work done during the thesis of Manolis Delakis which was developed in the frame of MUSCLE (see previous activity reports). The second chapter, written by Gwenole Lecorve, Stephane Huet, Pascale Sebillot, and Guillaume Gravier was devoted to the coupling between Naturel Language Processing and Automatic Speech Recognition. This chapter is based on some work done in the group independently of MUSCLE and opens new questions which will be part of our contribution until the end of the project. The main topic of this chapter is: How to improve speech recognition using NLP tools?

5.2 Contribution by UniS

Researchers involved

A Khan, W. Christmas

Activities & Achievements

Application of level set methods to lip tracking for multimodal speech understanding: We implemented, tested and modified some existing methods that integrate probabilistic information with level set methods for tracking. We then developed a new method that could use learned probabilistic information (obtained from support vector machine classification) and integrate it within an active contour/level set framework for segmentation and tracking of objects. We applied this technique for lip segmentation in human face images. We also worked on integrating probabilistic information with learned shape information for tracking in cluttered environment.

Publications

We wrote a paper on the new method described above that was accepted at the ISVC'07 (International Symposium on Visual Computing) conference.

5.3 Contribution by AUTH

Researchers involved

V. Moschou, P. Antonopoulos, E. Benetos, D. Ververidis, I. Kotsia, C. Kotropoulos, N. Nikolaidis, I. Pitas

Activities & Achievements

A new release (Release 1.2) of AUTH MUSCLE Movie Database has been prepared where corrections were made in order to improve synchronization between audio and video channels and the associated annotation. On-going research on speaker diarization within the showcase "Movie Summarization and Skimming Demonstration" has been continued aiming at dialogue detection in movie scenes. The overall classification error has been reduced to 12.56%, when outliers are not included in the clustering. However, the algorithm tends to over-estimate the number of clusters. Research is focused on finding a number of clusters that is close to ground truth. Three speaker segmentation systems have been assessed. The first system investigates the AudioSpectrumCentroid and the AudioWaveformEnvelope features, implements a dynamic fusion scheme, and applies the Bayesian Information Criterion (BIC). The second system consists of three modules. In the first module, a second-order statistic-measure is extracted; the Euclidean distance and the T2 Hotelling statistic are applied sequentially in the second module; and BIC is utilized in the third module. The third system first uses a metric-based approach, in order to detect potential speaker change points, and then the BIC criterion is applied to validate the previously detected change points. Experiments are carried out on a dataset, which is created by concatenating speakers from the TIMIT database. A systematic performance comparison among the three systems is carried out by means of one-way ANOVA method and post hoc Tukey's method.

Events

C. Kotropoulos attended 2007 IEEE Workshop Engineering Applications of Neural Networks held at Thessaloniki in August 2007.

Publications

V. Moschou, M. Kotti, E. Benetos, and C. Kotropoulos, "Systematic comparison of BIC-based speaker segmentation systems," in Proc. 2007 IEEE Multimedia Signal Processing Conf., accepted.

5.4 Contribution by TU Vienna - IFS

Researchers involved

Robert Neumayer, Andreas Rauber

Activities & Achievements

We prepared a bi-modal test collection for multi-modal audio experiments. Both audio features and song lyrics (fetched from the Internet), were extracted/computed and the experiment results were used for the Book Chapter in the MUSCLE Book on Multimodal Processing and Interaction, edited by Petros Maragos. The Book Chapter "Multi-Modal Analysis of Text and Audio Features for Music Information Retrieval" was finalized.

5.5 Contribution by ICCS-NTUA

Researchers involved

G. Papandreou, A. Katsamanis, V. Pitsikalis, P. Maragos

Activities & Achievements

Audio-Visual Interaction for Speech Processing
In our on-going effort in the field of multimodal fusion for Automatic Audio-Visual Speech Recognition and Processing, we have extended our approach to automatic multimodal fusion by uncertainty compensation by incorporating the effect of measurement uncertainty into model training procedures as well. This allows model training from noisy data and improves speech recognition performance, especially at slow SNR values.

Publications

G. Papandreou, A. Katsamanis, V. Pitsikalis, and P. Maragos, "Multimodal Fusion and Learning with Uncertain Features Applied to Audiovisual Speech Recognition", in Proc. IEEE Workshop on Multimedia Signal Processing (MMSP-2007), 2007.

5.6 Contribution by ICCS-NTUA

Researchers involved

ICCS-NTUA (G. Papandreou, P. Maragos) TSI-TUC (M. Perakakis, A. Potamianos, E. Sanchez-Soto) INRIA-Texmex (G. Gravier, P. Gros)

Activities & Achievements

Real-time Audio-visual Automatic Speech Recognition Demonstrator Showcase
We have continued during the reporting period the effort to build a real-time audio-visual automatic speech recognition demonstrator, as part of the Muscle Showcasing initiative. We have spent significant effort in streamlining the visual front-end of the prototype. Trying to overcome the limitations of our previous research-level visual front-end module, which was designed for off-line feature extraction in well-defined scenarios, such as processing videos shot under carefully controlled conditions, we have made the following improvements: We have integrated a fast adaboost-based face detector in our visual front-end. The role of the face detector is to initialize the face tracker at the first frame, and also whenever tracking fails. We have made algorithmic advances in the techniques for AAM fitting, yielding efficient and accurate computer vision algorithms for parametric model fitting and tracking. Repeated image resampling at updated shape warps (texture mapping in computer graphics

terminology) constitutes a significant part of the computational load during model fitting. We build on advanced features of modern graphics cards (GPUs) and perform texture-mapping very efficiently on the GPU. We are currently working on integrating all these components into our prototype's visual front-end.

5.7 Contribution by ICCS-NTUA

Researchers involved

G. Evangelopoulos, K. Rapantzikos, and P. Maragos

Activities & Achievements

Audiovisual Attention Modeling and Salient Event Detection
Although human perception appears to be automatic and unconscious there exist complex sensory mechanisms that form the preattentive component of human understanding and lead to awareness. Considerable research has been carried out into these preattentive mechanisms and computational models have been developed and employed to common computer vision or speech analysis problems. The separate audio and visual modules may convey explicit, complementary or mutually exclusive information around structures of audiovisual events. In any video sequence the two streams are processed in parallel. Based on recent studies on perceptual and computer attention modeling, we extract attention curves using features around the spatiotemporal structure of video and sounds. The potential of intra-module fusion and audiovisual event detection is demonstrated in applications such as key-frame selection, video skimming and summarization and audio/visual segmentation. During the reported period, there has been considerable progress in building an image saliency Matlab toolbox which will soon be put on the public domain. More specifically, this code is an implementation of T. Lindeberg's scale-invariant primal sketch used for object detection and shape analysis, and detects through an image scale-space salient edge and ridge contours. These contours are then post-processed to obtain a sparse set of straight line segments, which can then be utilized for higher level recognition tasks.

Publications

K. Rapantzikos, Y.Avrithis, S. Kollias, "Spatiotemporal saliency for event detection and representation in the 3D Wavelet Domain: Potential in human action recognition", in Proc. of ACM International Conference on Image and Video Retrieval (CIVR), Amsterdam, July 2007
K. Rapantzikos, G. Evangelopoulos, P. Maragos, and Y. Avrithis, "An audiovisual saliency model for movie summarization", in Proc. IEEE Workshop on Multimedia Signal Processing (MMSP-2007), 2007.

5.8 Contribution by ICCS-NTUA

Researchers involved

ICCS-NTUA (P. Maragos, G. Evangelopoulos, K. Rapantzikos, I. Avrithis) AUTH (C. Kotropoulos, P. Antonopoulos, V. Moschou, N. Nikolaidis, I. Pitas) INRIA-Texmex (P. Gros, X. Naturel) TSI-TUC (A. Potamianos, E. Petrakis, M. Perakakis, M. Toutoudakis)

Activities & Achievements

Movie Summarization and Skimming Demonstrator Showcase
During the reporting period, we have launched an effort to build a Movie Summarization and Skimming Demonstrator, as part of the Muscle Showcasing initiative Participating partners are ICCS-NTUA (leader), TSI-TUC, AUTH, and INRIA-Texmex. As the amount of video data available (movie, TV programs, clips) in a personal recorder or computer are becoming increasingly large (100h in VCRs or hundreds of hours on a PC) intelligent algorithms for efficiently representing video data and presenting them to the user are becoming important. Video summarization, movie summarization and movie skimming are increasingly popular research areas with immediate applications. In this showcasing project we will: (i) use combined audio and video saliency detectors to identify the importance of movie content to the user and (ii) design an interface that presents the audio and video information to the user in a compressed form, thus saving time with little or no loss of information. The demonstrator will have the ability to render a movie from its typical 2h duration down to 30' by skimming over (fast forwarding or omitting) non-salient movie scenes while playback at regular speed parts of the movie with salient audio and video information. The interface will also have the ability to break the synchrony of the audio/video streams and selectively present audio or video information.

Events

Regarding dissemination activities, the demonstrator was showcased as a Technical Demo at the ACM International Conference on Image and Video Retrieval (CIVR), held in Amsterdam on July 9-11 2007.

5.9 Contribution by ICCS-NTUA

Researchers involved

Petros Maragos (ICCS-NTUA) Alexandros Potamianos (TSI-TUC) Patrick Gros (INRIA-Texmex)

Activities & Achievements

Book on "Multimodal Processing and Interaction: Audio, Video, Text"
Petros Maragos (ICCS-NTUA), Alexandros Potamianos (TSI-TUC) and Patrick Gros (INRIA-Texmex), Editors.
The book under preparation is covering the thematic areas of WP5 (former JPA2 WP6 & WP10). It comprises two main parts: Part A will be a comprehensive State-of-the-Art review of the area and Part B will consist of selected research contributions / chapters by Muscle WP5 members. The book proposal was submitted to Springer-Verlag and has been approved.

5.10 Contribution by ICCS-NTUA

Researchers involved

A. Katsamanis, A. Roussos, G. Papandreou, P. Maragos (ICCS-NTUA) Y. Laprie (INRIA-LORIA)

Activities & Achievements

Audio-Visual Speech Inversion
In this research track we address the problem of audiovisual speech inversion, namely recovering the vocal tract's geometry from auditory and visual speech cues. In our recent work, which has culminated into a submitted paper during the reported period, we approach the problem in a statistical framework, combining ideas from multistream Hidden Markov Models and canonical correlation analysis, and demonstrate effective estimation of the trajectories followed by certain points of interest in the speech production system. Our experiments show that exploiting both audio and visual modalities clearly improves performance relative to either audio-only or visual-only estimation. We report experiments on the QSMT database which contains audio, video, and electromagnetic articulography data recorded in parallel.

Publications

A. Katsamanis, G. Papandreou, and P. Maragos, "Audiovisual-to-Articulatory Inversion Using Hidden Markov Models", in Proc. IEEE Workshop on Multimedia Signal Processing (MMSP-2007), 2007.

6 Overview activities in WP6

6.1 Contribution by ISTI-CNR

Researchers involved

Luigi Bedini, Sara Colantonio, Ercan Kuruoglu, Massimo Martinelli, Davide Moroni, Emanuele Salerno, Ovidio Salvetti

Activities & Achievements

Within the activities of e-team 13, we are continuing the experimentation on the MCMC technique for blind separation, the evaluation of the results from a Fourier-domain version of the Correlated Component Analysis, and the development of the MaxNG method applied to hyperspectral remote-sensed data. Collaborations are active both within the MUSCLE network (TCD) and other external partners (Padova Astronomical Observatory, Italy, Complex Systems Laboratory University of Buenos Aires, Argentina, IIA-LARA-CNR, Italy). A theoretical-experimental paper on the application of MaxNG to remote sensing has

been accepted for publication in Signal Processing, and published online by the ScienceDirect system. An e-team meeting has been held in Pisa from 23 to 27 July.

The activity on the development of a hybrid approach for supporting decision making has been carried on addressing the problem of how integrate inferential and computational reasoning through meta-level reasoning capabilities. Within the frame of decision making based on image understanding, an activity has started for evaluating the application of the Algorithm ontology. The results obtained from the definition of ontology for supporting the feature-based analysis of microscopic cell images were presented at the 7th Open German/Russian Workshop, held in Ettlingen – Germany on August 20-23.

Events

Simon Wilson gave a talk at ISTI on 27 July. Title: Fully Bayesian Source Separation with Application to the Cosmic Microwave Background.

Participation to the 7th Open German/Russian Workshop Pattern Recognition and Image Understanding, August 20-23, Ettlingen, Germany

Publications

C. F. Caiafa, E. Salerno, A. N. Proto, L. Fiumi, "Blind spectral unmixing by local maximization of non-Gaussianity", Signal Processing, Vol. 88, No. 1, pp. 50-68, January 2008, doi:10.1016/j.sigpro.2007.07.011

S. Colantonio, I.B. Gurevich, M. Martinelli, O. Salvetti, Yu. Trusova, "Ontology driven Approach to Cell Image Analysis", in Open German Russian Workshop on Pattern Recognition & Image Understanding (Ettlingen, Germany, 20-23 August 2007), Forshungsinstitut fur Optronik und Mustererkennung, 2007.

6.2 Contribution by TCD

Researchers involved

Simon Wilson Ercan Kuruoglu (CNR-ISTI) Emanuele Salerno (CNR-ISTI)

Activities & Achievements

Simon Wilson visited ISTI in Pisa for 1 week as part of the planned activities of E-team 14 "Unsupervised segmentation and classification of multichannel data". Data on extraterrestrial microwave radiation from the WMAP satellite was downloaded and prepared for use as a blind source separation problem. The Bayesian inference algorithm used to do the separation was agreed upon. The outline of a paper to be submitted to an upcoming edition of IEEE JSAC on astronomical data was prepared.

6.3 Contribution by TUG

Researchers involved

Thomas Mauthner, Amir R. Saffari A. A., Horst Bischof

Activities & Achievements

1. We did further extensions on our probabilistic tracking framework. Additional cues to distinguish several similar looking object during tracking were implemented and tested. Due to this enhancement the robustness of the tracker is considerably improved. 2. We are currently continuing the experiments on large-scale visual object recognition problems where the number of features to be clustered are very large for classical algorithms (e.g. 1000000 SIFT descriptors). Our methods show very interesting results and achieve relatively good performance in a reasonable time.

6.4 Contribution by UCD

Researchers involved

Pádraig Cunningham Derek Greene Ken Bryan

Activities & Achievements

The penultimate draft of the Springer edited volume on Machine Learning Techniques for Processing Multimedia Content was produced at the end of August. At present some minor issues associated with colour and generation of the index need to be resolved and then the source can be handed over to Springer. At present we have considerable activity on One-Class Classification (OCC) techniques in UCD. We are working on the use of OCC for speaker verification and also on the problem of dimension reduction for OCC. In August we published two papers at the Irish AI conference presenting some preliminary results of this work.

Events

Pádraig Cunningham, Derek Greene & Ken Bryan attended the Irish AI conference at the end of August 20007.

Publications

Villalba, S., Cunningham, P., (2007) An Evaluation of Dimension Reduction Techniques for One-Class Classification, Proceedings of 18th Irish Conference on Artificial Intelligence and Cognitive Science, S. J. Delany, M. Madden (eds). Brew, A., Grimaldi, M., Cunningham, P., (2007) An Evaluation of One-Class Classification Techniques for Speaker Verification, Proceedings of 18th Irish Conference on Artificial Intelligence and Cognitive Science, S. J. Delany, M. Madden (eds).

6.5 Contribution by INRIA Ariana

Researchers involved

Josiane Zerubia, Ian Jermyn, Ting Peng.

Activities & Achievements

Ting Peng, joint PhD student of INRIA Ariana and the LIAMA Institute in Beijing, has created a phase field version of the new linear model described under WP3, and has implemented an efficient gradient descent algorithm for this model. INRIA Ariana has a INRIA Associated Team project 'Shapes' with Professors Anuj Srivastava and Victor Patrangenaru of the Statistics Department of Florida State University As part of this collaboration, Ian Jermyn has been working with Professor Srivastava on new Riemannian metrics on shape spaces derived from the Fisher-Rao metric on the space of (probability) measures. These new metrics, under which the space of parameterized curves in any number of dimensions becomes Euclidean, greatly simplify the calculation of geodesics and other quantities of geometric and statistical significance. They are currently applying the theory to the extraction of shapes from noisy point clouds.

Events

CVPR 2007, EMMCVPR 2007.

Publications

Two papers have been published in CVPR 2007, and another in EMMCVPR 2007, on the research resulting from the collaboration with Professor Srivastava.

6.6 Contribution by ENSEA

Researchers involved

Prof. Sylvie Philipp-Foliguet, Philippe-Henri Gosselin, Justine Lebrun.

Activities & Achievements

We continued our works on interactive research in database of images containing a type of object or person. Images are represented by an adjacency graph of fuzzy regions, with their colour and texture descriptors. The problem is the inexact match of graphs or subgraphs, which is studied through kernels on paths of these graphs. We worked on a general framework using various kernels and various sets of paths and obtained interesting results both in terms of precision and of execution time, notably on the Caltech database.

Publications

This work is going to be submitted at first in a French conference: EGC 2008, Sophia Antipolis, France.

6.7 Contribution by UPMC

Researchers involved

Matthieu Cord, David Picard, Julien Gony

Activities & Achievements

We had activities for the period on Image indexing and Machine Learning based on reinforcement learning techniques for the E-Team "Active and Semi-Supervised Learning e-team" and Information retrieval for multimedia documents (text+images). We made experiments on TrecVid corpus for reinforcement learning based on distributed databases over a network of computers. We also continued the collaboration with F. Precioso (ENSEA) on developing a visual semantic class learning system for video content analysis. We prepared oral presentations (with PH. Gosselin and S. Philipp-Foliguet (ENSEA)) on "Kernel on Bags of Fuzzy Regions for fast object retrieval" and "Descriptor matching for image identification on cultural databases" for the IEEE International Conference on Image Processing 2007, that takes place in San Antonio, Texas, USA in September 2007. Matthieu Cord still worked on Editing the collective MUSCLE book on ML for Multimedia to be published by Springer. We participated to the Fifth International Workshop on Content-Based Multimedia Indexing (CBMI 2007) with special MUSCLE event. We also participate to the ACM International Conference on Image and Video Retrieval, July 9-11 2007, at University of Amsterdam, The Netherlands. In this important event for the multimedia retrieval community, we presented one paper during the poster session. We also had a demo of the RETIN system (developed with the ENSEA partner) accepted and presented during the demo session. Additionally, we presented the system during the special MUSCLE session organized by our network coordinator N. Boujemaa. Finally, we also participated to the image Retrieval Showcase coordinated by the MUSCLE member Allan Hanbury, Vienna University of Technology, Austria. We won the price of the best recall during this online competition.

6.8 Contribution by TU Vienna - IFS

Researchers involved

Robert Neumayer, Khalid Latif, Rudolf Mayer, Andreas Rauber

Activities & Achievements

The final version of the paper "Sky-Metaphor Visualisation for Self-Organising Maps", accepted for publication at the 7th International Conference on Knowledge Management (I-KNOW'07), Graz, Austria, September 5 - 7 2007, was prepared. A poster was made for presentation of our work on the metro visualisation of component planes for self-organising maps at the International Joint Conference on Neural Networks (IJCNN'07), Orlando, FL, USA, August 12 - 17 2007.

Events

International Joint Conference on Neural Networks (IJCNN'07), Orlando, FL, USA, August 12 - 17 2007. International Conference on Knowledge Management (I-KNOW'07), Graz, Austria, September 5 - 7 2007.

Publications

Robert Neumayer, Rudolf Mayer, Georg Pölzlbauer, and Andreas Rauber. The metro visualisation of component planes for self-organising maps. In Proceedings of the International Joint Conference on Neural Networks (IJCNN'07), Orlando, FL, USA, August 12 - 17 2007. IEEE Computer Society.

Khalid Latif and Rudolf Mayer. Sky-Metaphor Visualisation for Self-Organising Maps (accepted for publication). In Proceedings of the 7th International Conference on Knowledge Management (I-KNOW'07), Graz, Austria, September 5 - 7 2007

6.9 Contribution by AUTH

Researchers involved

C. Kotropoulos

Activities & Achievements

Adaptive genetic algorithms are employed to search for the worst performing features with respect to the probability of correct classification achieved by the Bayes classifier in a first stage. These features are subsequently excluded from sequential floating feature selection that employs the probability of correct classification of the Bayes classifier as criterion. In a second stage, adaptive genetic algorithms search for the worst performing utterances with respect to the same criterion. The sequential application of both stages is demonstrated to improve speech emotion recognition on the Danish Emotional Speech database.

Events

C. Kotropoulos attended 2007 IEEE Workshop on Machine Learning for Signal Processing held at Thessaloniki in August 2007.

6.10 Contribution by IBaI

Researchers involved

Petra Perner Horst Perner Suzanne Little

Activities & Achievements

The prototype-based classifier has been set up for the needs within MUSCLE and tested against various data bases. Studies on Feature Weighting and Feature Selection have been done. The dependency between Feature Weighting and Feature Selection and Prototype

Selection on various Data bases from UCI Repository has been studied. The results got reported and new development issues have been drawn from these results.

7 Overview activities in WP7

7.1 Contribution by TAU-VISUAL

Researchers involved

Tomer AMIAZ and Nahum KIRYATI. Collaboration with MTA-SZTAKI.

Activities & Achievements

Towards meeting the challenge of the dynamic texture segmentation showcase, we focused on the development, testing and documentation of two approaches for fast dynamic texture segmentation. Taking our variational approach as a starting point, two approaches were developed and tested. One is a graph-cut based implementation, the other is a simplified variational method. Experimental results indicate that the simplified variational method leads to good results, close to the results yielded by the full variational model. We have not been able yet to obtain similarly good results with the level-set based approach.

7.2 Contribution by Bilkent University

Researchers involved

Kivanc Kose, Erdem Dengel, A. Enis Cetin, Behcet Ugur Toreyin, Yigithan Dedeoglu, Ugur Gudukbay

Activities & Achievements

The preparations for IBC'07 exhibition are continued. The posters and the videos which will be presented at the exhibition are prepared. Since MUSCLE will have 8 demos at the exhibition, a time schedule for the demos is prepared. All the partners are informed about this schedule, and what kind of dissemination materials they need to bring with.

Contracts for booths at CEBIT Eurasia'07 and INTERSEC'08 exhibitions are signed. MUSCLE showcases will be presented at those exhibitions. We are planning to collect all the materials which will be exhibited at IBC'07, and present them at CEBIT Eurasia and INTERSEC.

Events

IBC'07 exhibition will be held on 7-11 September 2007 in Amsterdam. CEBIT Eurasia'07 exhibition will be held on 2-7 October 2007 in Istanbul. INTERSEC'07 exhibition will be held on 13-15 January 2008 in Dubai.

7.3 Contribution by Bilkent University

Researchers involved

Muhammet Bastan, Onur Kucuktunc, Hayati Cam, Ugur Gudukbay, and Ozgur Ulusoy

Activities & Achievements

We continued our research on integration of structural and semantic models for multimedia metadata management. Within this context, we implemented automatic object extraction and tracking for our video database management system. We used Open CV and FFMPEG libraries for this purpose. We start to work on developing a tool for automatic extraction of MPEG-7 features (color, texture, etc.) to be used for querying in video databases. We plan to demonstrate the tool as a showcase in the e-team workshops to be held in the following months.

7.4 Contribution by UPC

Researchers involved

Montse Pardas, Dafnis Batalle

Activities & Achievements

Contribution to the Showcase "Unusual event detection" . An events detection module based on data produced by the body and motion analysis has been developed. A vector of features is extracted for classification, including axis of the ellipse fitted to the detected foreground object, orientation of this ellipse, and Hu moments of the Motion History Image and Motion Energy Image. During this period the tools developed at UPC for Human model and motion based unusual event detection have been adapted to recognize different types of events in a controlled entrance. The types which are distinguished are: normal, jumping and passing below.

7.5 Contribution by GET-ENST

Researchers involved

Beatrice Pesquet-Popescu, Maria Trocan

Activities & Achievements

We pursued the work on video analysis for fire detection, based on motion-compensated multiresolution analysis. Progress has been made in the application of graph-cut method. In particular, the modelization of spatio-temporal subbands and even at lower level, for blocks in these subbands, and of their relations was done in order to allow the application of this energy-minimization method. Several papers have been submitted to conferences. Another idea that we would like to develop was the synthesis of fire textures through statistical interpolation of small patches of such textures. A post-doc was found to work on this topic.

7.6 Contribution by TU VIENNA-PRIP

Researchers involved

Julian Stöttinger, Allan Hanbury

Activities & Achievements

The object recognition showcase system was demonstrated at the demo session of the CIVR on the 9th and 11th of July 2007. Further work was done to improve the system. This involved improving the classifier and improving the repeatability of the colour interest points (interest points with good repeatability should appear in the same place on objects even when the objects are seen from a different viewing angle). For classifier improvement, support vector machines (SVMs) were incorporated into the system. The improvement in repeatability was brought about by testing a range of parameters to find those which produce the highest repeatability.

Events

Demonstrated at the demo session of the CIVR on the 9th and 11th of July 2007.

Publications

J. Stöttinger, A. Hanbury, N. Sebe and T. Gevers, "Do Colour Interest Points Improve Image Retrieval?", Proceedings of the 14th IEEE International Conference on Image Processing (ICIP'07), San Antonio, Texas, September 16-19, 2007

7.7 Contribution by IRIT-UPS

Researchers involved

P. Joly, J. Pinquier, F. Giani

Activities & Achievements

The extension for JPA4 of the showcase ACADI (Automatic Character in Audiovisual Document Indexing) was accepted by the Steering Committee. During CIVR'07 which was held in Amsterdam (9-11 July) was presented the functionalities of the web portal developed for the CASEWP showcase. The demonstrated functionalities were : - the access to a content set, - the result submission, - the result parsing, - the result evaluation, - the synthetic graphics generation with different possible options.

Events

Participation to ACM International Conference on Image and Video Retrieval – CIVR'2007 in Amsterdam (The Netherlands) on July 9-11 2007: presentation of the ACADI showcase and the CASEWP showcase.

7.8 Contribution by INRIA Ariana

Researchers involved

Josiane Zerubia, Ian Jermyn, Peter Horvath.

Activities & Achievements

In July, Ian Jermyn and Josiane Zerubia attended a meeting of the ACI QuerySat project, in which the Ariana project-team is a partner, and presented MUSCLE-related work. Ian Jermyn and Josiane Zerubia met with representatives of EADS MBDA, and discussed possible joint projects. This will lead to the submission of an EADS Foundation proposal for funding. Ian Jermyn presented MUSCLE-related work to representatives of CS. In August, Ian Jermyn attended CAIP 2007 and Peter Horvath attended ACIVS 2007, to present MUSCLE-related work.

7.9 Contribution by TU Vienna - IFS

Researchers involved

Thomas Lidy, Andreas Rauber

Activities & Achievements

The Content Analysis Showcase and Evaluation Web Portal was shown at ACM International Conference on Image and Video Retrieval, July 9-11 2007, Amsterdam, The Netherlands. A showcase meeting took place also there. Preparations for lectures at the First Russian Summer School in Information Retrieval were done. For the MUSCLE stand at IBC (International Broadcasting Convention) in Amsterdam several posters were prepared, flyers were designed and printed PowerPoint slides and a live demonstration were prepared.

Events

ACM International Conference on Image and Video Retrieval, July 9-11 2007, Amsterdam, The Netherlands. First Russian Summer School in Information Retrieval, September 5-12, 2007, Ekaterinburg, Russia. IBC - International Broadcasting Convention, September 7-11, 2007, Amsterdam, The Netherlands.

7.10 Contribution by IBaI

Researchers involved

Petra Perner Horst Perner Suzanne Little

Activities & Achievements

The recent tool on prototype-based classification has been demonstrated to various industrial and industrial-related research institutes in Germany. The application issues for this method

have been explained to them as well as the advantages for the proposed application. This has been done for various biotechnological and medical applications. The method has also been demonstrated in the Industrial Tutorial on Data Mining.

8 MUSCLE Effort Table

Person-Months report for period : July - August 2007

Institute	WP1a	WP1b	WP2	WP3	WP4	WP5	WP6	WP7	Total
01 - ERCIM	1.2	0	0	0	0	0	0	0	1.2
03 - UCL	0	0	0	2	0	0	0	0	2
04 - KTH	0	0	0	0	0.2	0.7	0	0	0.9
05 - BILKENT	0	0.2	0.3	0	0.25	0.4	0.4	4.5	6.05
06 - VIENNA PRIP	0	0	0.2	1.4	0	0	0	1.5	3.1
07 -MTA SZTAKI	0	0	0	0.8	0	0	1.5	1.7	4
09 - CNR-ISTI	0	0.32	0.64	0.32	0	0	0.32	0	1.6
11 - TUG	0	0.004	0	1.96	0	0	1.89	0.17	4.024
12 - UPC	0	0	1	0.2	0	0	0	0.5	1.7
14 - UTIA	0.1	0	0	1	0.5	0.1	0.3	0.5	2.5
16 - AUTH	0	0	0	0	0	3.521	0	0	3.521
17 - CEA	0	0.1	0.5	0.1	0	0	0.2	0	0.9
18 - TU VIENNA IFS	0	0	0	0	2.5	0.25	1	1.5	5.25
19 - ACV	0	0	0	0.4	0	0	0	0	0.4
22 - IBAI	0	0.01	0.33	0	0	0	0.76	0.33	1.44
26 - TAU-SPEECH	0	0	0	0	4	0	0.3	0	4.3
27-TAU-VISUAL	0	0	0	2	0	0	0	1	3
32 - INRIA Ariana	0	0.18	0	0.38	0	0	0.35	0.32	1.23
32- INRIA Tex Mex	0	0	0	0	0	1	0	0	1
32 - INRIA Vista	0	0.2	0	2.2	0	0	0	0	2.4
33 - GET	0.1	0	0	1.34	0	0	0	2	3.44
35 - UNIS	0	0	0	1	0	1	0	0	2
37 - ENSEA	0	0	0	0.35	0	0	0.35	0	0.7
39 - UPS – IRIT	0	0	0	1	1	0	0	1.5	3.5
41 - UPMC	0	0	0	0	0	0	1	0	1
42 -NUID / UCD	0	0.1	0	0	0	0	2.1	0	2.2
Total	**1.4**	**1.114**	**2.973**	**16.45**	**8.45**	**6.971**	**10.478**	**15.52**	**63.355**