MUSCLE

Network of Excellence

Multimedia Understanding through Semantics, Computation and Learning

Project no. FP6-507752

Bi-monthly progress report # 17: Nov-Dec 2006

Due date of deliverable: 15.01.2007
Actual submission date: 15.04.2007

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

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

Draft 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

- Administrative and financial coordination of the network
- Organisation of regular audio-conferences.
- E-Teams re-organisation
- Launch of Showcase project call
- Preparation and Organisation of the MUSCLE Plenary meeting.
- Migration of MUSCLE website content from CWI to ERCIM and re-design of the website
- Reimbursement of MUSCLE integration expenses (mobility support grant)

Events

MUSCLE Plenary meeting took place on 6, 7, 8 December 2006 in Paris. This meeting gathered 90 persons. During the first day, coordination issues were presented. A global vision of JPA3, the new WP structure, integration instrument: e-teams, showcases, and connections with other European projects were addressed by Nozha Boujemaa. In addition, each WP leader gave a WP overview. On the second day, the 15 e-Teams gave a presentation of their activity and progress. Some showcase proposals were presented as well. A second showcase proposal presentation session took place on the third meeting day, followed by workpackage scientific presentations. The steering Committee met to select the showcase proposals to be funded and selected 12 projects over 17 proposals. The agenda and the photos of the plenary meeting are available on the website:
Agenda: http://www.muscle-noe.org/content/view/89/40/
Photos: http://www.muscle-noe.org/content/view/136/37/

2 Overview activities in WP2

2.1 Contribution by partner CEA

Researchers involved

Moëllic Pierre-Alain

Activities

Evaluation Campaign ImagEVAL 2006, http://www.imageval.org/e_resultats.html

Achievements

End of the campaign. Analyse of the results and presentation of the results during the NicephoreDays 2006. Organization of the ending scientific workshop

Events

NicephoreDays 2006 (http://www.nicephoredays2006.com/en/2.htm) ImagEVAL 2006 Workshop (12/07/2007 with CIVR2007)

2.2 Contribution by partner IRIT-UPS

Researchers involved

Philippe Joly, Jerome Louradour, Jerome Farinas, Julien Pinquier

Activities

IRIT researchers start their activities in the framework of the MUSCLE NOE at the beginning of this period. These activities are related with: developing benchmarks and organizing evaluation campaign on video content indexing, participating to international evaluation campaign - distributed strategies for multimedia analysis tools integration.

Events

Participation to the Paris Meeting proposition of a showcase (TSEWP) on online temporal segmentation evaluation. This project has been merged with the "CAS" proposition to become the "CASEWP" (N°9 - accepted)

2.3 Contribution by partner TU VIENNA-PRIP

Researchers involved

Allan Hanbury, Branislav Micusik

Activities

1) Preparation of the evaluation showcase proposal (accepted in December)
2) Wrote articles on the coin competition 2006 results for the IAPR newsletter and the ERCIM News. Started preparing the 2007 coin competition
3) Updated the animal database annotations for the E-team on "Choosing Features for CBIR and Image Annotation"

Achievements

The evaluation showcase proposal was selected for funding

Problems

The website has been partly updated with new WP2 data. Finalizing the update for the deliverable has still been delayed so as to achieve better harmonization with the new MUSCLE main webpage. As the remaining changes mostly involve formatting and not content, this is not seen as a major problem

Events

Presentations on workpackage activities and plans at the MUSCLE general meeting in Paris.
Attended the ImagEVAL workshop in Chalon Sur Saône, France. Introduced the MUSCLE Evaluation Showcase in a short presentation

2.4 Contribution by partner ENSEA

Researchers involved

Sylvie Philipp-Foliguet, Philippe-Henri Gosselin, Frédéric Precioso, Julien Gony, Matthieu Cord (LIP6 Partner)

Activities

End of the campaign. Analyze of the results and presentation of the results during the NicephoreDays 2006

Events

NicephoreDays 2006 (http://www.nicephoredays2006.com/en/2.htm) ImagEVAL 2006 Workshop (12/07/2007 with CIVR2007) Participation to the final meeting of the Campaign at Chalon sur Saône 11-12-13-14 December 2006

Other

This Evaluation Track is organized by CEA Institute, one of the MUSCLE Members

2.5 Contribution by partner CNR-ISTI

Researchers involved

Suzanne Little, Massimo Martinelli, Patrizia Asirelli, Marco Tampucci, Ovidio Salvetti

Activities

A presentation about the MUSCLE Fellowship, titled "Multimedia metadata - indexing and integration using semantic web technologies", was prepared and presented at the MUSCLE Plenary held in Paris from Dec 6 to Dec 8. A proposal was prepared for a MUSCLE E-Team titled "Integration of structural and semantic models for multimedia metadata management" and a presentation also made at the plenary. Preliminary work on the E-Team website was started after the plenary. The paper titled "Multimedia Metadata Management: a proposal for an Infrastructure" (Asirelli, Little, Martinelli, Salvetti) was accepted to the SWAP workshop held in Pisa from Dec 18-20 and a presentation was given. Participation to W3C MMSEM XG has continued.

Achievements

The software 4M has been improved by adding few additional functionalities to the database management and the user interface

Events

MUSCLE Plenary held in Paris - W3C MMSEM face-to-face meeting in Athens - SWAP Conference, Pisa

Publications

"Multimedia Metadata Management: a proposal for an Infrastructure" (Asirelli, Little, Martinelli, Salvetti), Proc. SWAP, Pisa, 2006.

3 Overview activities in WP3

3.1 Contribution by partner CEA

Researchers involved

Moëllic Pierre-Alain, Millet Christophe, Hede Patrick

Activities

Features extraction and testing for the E-Team "Choosing Features for CBIR and Automated Image Annotation"

Achievements

Testing different features extraction techniques mainly based on colour and texture for animal detection

3.2 Contribution by partner UCL

Researchers involved

Fred Stentiford, Shijie Zhang

Activities

1) Experiments using attention based architectures for motion detection and estimation have shown that object motion in some circumstances may be estimated more accurately than other methods. Publications are in preparation
2) A paper and demonstration of attention based focusing are planned for the CIVR conference in July 2007
3) Further exchanges have taken place between INRIA and UCL continuing the collaboration on copy detection. A joint paper describing objective performance comparisons has been submitted to CIVR
4) Attended Muscle plenary meeting in Paris 6-9 Dec

Achievements

A proposal for a workshop entitled Computational Attention and Applications has been accepted for the International Conference on Computer Vision Systems to be held in Bielefeld in March 2007. A call for papers has been issued. Muscle sponsorship and a keynote speaker will be sought

3.3 Contribution by partner TAU SPEECH

Researchers involved

Arie Yeredor

Activities

Work on blind separation of superimposed images with relative shifts / cross-fade mixed video sequences, enabling subsequent processing of each separated source

Achievements

Successful separation of superimposed images with relative spatial shifts (demonstrated at www.eng.tau.ac.il/~arie/imsep.htm), including video sequences

3.4 Contribution by partner TAU-VISUAL

Researchers involved

Tammy Riklin-Raviv, Leah Bar, Tomer Amiaz, Nir Sochen, Nahum Kiryati

Activities

Preparation and submission of a journal article: "Shape-based Mutual Segmentation". Preparation and submission of a paper to the ISBI'07 conference: "Propagating Distributions for Segmentation of Brain Atlas" (together with researchers from Drexel University)

3.5 Contribution by partner ISTI-CNR

Researchers involved

Anna Tonazzini, Luigi Bedini

Activities

Study on blind separation methods for the clustering of gene expression profiles from DNA microarray experiments. Dissemination of the results of past research and technology transfer activity on document analysis

Events

Activities in document image processing presented at: International Conference Museums, Libraries and Archives Online: MICHAEL service and other international initiatives (4-5 December 2006, Roma)

Publications

- E. Console, V. Burdin, G. Cazuguel, S. Legnaioli, V. Palleschi, R. Tassone, A. Tonazzini, “Virtual restoring by multispectral imaging”, International Conference Museums, libraries and archives online: MICHAEL service and other international initiatives (4-5 December 2006, Roma)
- Poster. E. Console, V. Burdin, G. Cazuguel, S. Legnaioli, V. Palleschi, R. Tassone, A. Tonazzini, "Isyreadet: un sistema integrato per il restauro virtuale di documenti", Annali della Facoltà di Economia (n.12, 2006), Università del Sannio

3.6 Contribution by partner SZTAKI

Researchers involved

Dmitry Chetverikov, Sandor Fazekas

Activities

Developing methods for automatic detection, segmentation and recognition of dynamic texture in video

Achievements

A paper on DT detection submitted to SSVM 2007 jointly with TAU-Visual. Another related paper submitted to the Hungarian Conference on Image Processing

Events

Chetverikov and Fazekas participated in the plenary meeting in Paris where the program for DT detection was presented

3.7 Contribution by partner UPC

Researchers involved

Montse Pardas, Xavier Giró, Javier Ruiz, Camilo Dorea, Cristian Canton

Activities

1) Computation of region features for the e-team "Choosing features for CBIR"

2) Creation of demo sequences for a new representation tool for video sequences that allows description and indexation based on video objects

3) Organization of the Workshop of the e-team "Person detection, recognition and tracking"

4) Contribution to the joint submission of two papers to VISAPP, one of them with AUTH and the other with Bilkent

Events

Muscle Plenary meeting

3.8 Contribution by partner INRIA Ariana

Researchers involved

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

Activities

Ariana's work in WP3 involves three areas: modelling images, modelling regions in the image domain, and indexing. The first two, as likelihood and prior, enable the extraction of the regions in images corresponding to particular entities. The third uses this information to retrieve images from databases.

Image modelling: Ting Peng, joint PhD student of INRIA-Ariana and the LIAMA Institute in Beijing, has continued to work on image models appropriate for road network segmentation from very high resolution (0.5m) satellite images. The previous image model captured the one-point statistics of scaling coefficients through mixture of Gaussian models, while the two point statistics were modelled using the variance of the values in a window. The latter method takes no account of the spatial distance between the dependent variables, and seems to be one of the reasons why the segmentation results were not completely satisfactory. Her latest research uses Gaussian distributions with inverse covariance consisting of linear combinations of the Laplacian operator and a constant. This enables the modelling of the different levels of smoothness in the background and in the road. In this way, the two-point statistics are more effectively captured, while the model is simpler. The experimental results based on this model are very good, producing near-perfect segmentations of major roads at 1/8 resolution. The challenge remains how to make use of the finest scale information in the data, which has a very complex structure, and how to deal with widely varying road widths. Region modelling: Ting Peng has also been working on including existing Geographical Information System (GIS) data in the prior model.

This is particularly important for the map updating application of Ms. Peng's work, since such information frequently exists and can help significantly. Ms Peng has developed two energy terms for the inclusion of GIS data.

One in particular is capable of correcting both missing and extra roads in old GIS data based on information from the newer image data, while at the same time using the GIS data to improve the segmentation of roads that appear in both the GIS and image data. Region modelling: the higher-order active contour (HOAC) model for a 'gas of circles' has been

further advanced by Peter Horvath, joint PhD student of INRIA-Ariana and the University of Szeged. This model is being applied to the extraction of tree crowns from remote sensing images. Very good segmentation results are now being obtained with a minimum of parameters: some parameters are fixed using the stability calculations, while others are learned from examples. The few remaining parameters are still fixed by hand. Indexing: Avik Bhattacharya has continued his work on using road networks for retrieval from remote sensing image databases. Preliminary results on a small database with a limited number of classes are very promising. Work is now concentrating on using an expanded feature set, with feature selection, to classify a larger database into a finer set of classes.

Publications

- MP [544], MP [541], MP [540], MP [542]

3.9 Contribution by partner UTIA

Researchers involved

M.Haindl, P. Vacha

Activities

Illumination invariants for CBIR

Achievements

We proposed fast and robust image retrieval measures 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.

3.10 Contribution by partner TUG

Researchers involved

Martin Winter, Horst Bischof

Activities

We developed a framework to substantially increase the recognition performance of a vocabulary tree based recognition system. This is done by combination of hypothesis obtained by a standard inverse object voting algorithm with reliable spatial relations descriptors. The algorithm operates on different depths of a standard k-means tree, coevally benefiting from the advantages of different levels of information abstraction.

3.11 Contribution by partner IBaI

Researchers involved

Petra Perner, Horst Perner

Activities

We have developed a framework for learning semantic inference rules for images and videos

Achievements

The framework has been implemented into a program

Publications

- Image Mining for the Construction of Semantic-Inference Rules and for the Development of Automatic Image Diagnosis Systems, P. Perner In: Xingquan Zhu and Ian Davidson, Knowledge Discovery and Data Mining: Challenges and Realities with Real World Data, IDEA Group Inc. Publisher

3.12 Contribution by partner ARMINES

Researchers involved

B. Marcotegui & F. Bach

Activities

1/ Submission of a joint paper (A. Hanbury and B. Marcotegui) "Morphological segmentation on learned boundaries" to Image and Vision Computing
2/ Ongoing work on the development of a family of kernels between two segmentation graphs
3/ First experiments on texture features based on morphological filters (ultimate opening)

3.13 Contribution by partner ACV

Researchers involved

Csaba Beleznai, Herbert Ramoser

Activities

Preparations for Showcase 6 "Real-Time Detector For Unusual Behaviour" Participation in the plenary meeting

Events

MUSCLE Plenary Meeting Paris, Dec. 6-8, 2006

3.14 Contribution by partner IRIT-UPS

Researchers involved

Philippe Joly, Elie El Khoury

Activities

IRIT researchers start their activities in the framework of the MUSCLE NOE at the beginning of this period. These activities are related with person in video content identification.

3.15 Contribution by partner CNR-ISTI

Researchers involved

Davide Moroni, Sara Colantonio, Ovidio Salvetti

Activities

Activity performed in the e-team 4 has been presented at Muscle plenary meeting in Paris. Specific work has been focused on the automatic spatio-temporal analysis and recognition of deformable structures, from the points of view of assessment of the dynamic behavior of a structure and classification of its deformation pattern. The method developed was partially tested on the elective study case of heart dynamics.

Achievements
The prototypical software tool already developed has been improved by adding a new functionality.

Events
Meeting in Moscow with DCC dept. of RAS - Seminar "A General Approach to Shape Characterization for Biomedical Problems", Moscow, Nov. 2006

3.16 Contribution by partner UNIS

Researchers involved
W. Christmas, F. Yan, I.Kolonias

Activities
Preparing RAVL Vision and Pattern Recognition C++ library for release

Publications
- B Goswami, W Christmas, and J Kittler. Stastical estimators for use in automatic lip segmentation. In 3rd European Conference on Visual Media Production, pages 79-86, 2006
- J-Y Guillemaut, J Kittler, M Sadeghi, and W Christmas. General pose face recognition using frontal face model. In J Martinez-Trinidad, J Ochoa, and J Kittler, editors, Proceedings of the 11th Iberoamerican Congress in Pattern Recognition, pages 79-98. Springer, November 2006
- Hory, W Christmas, and A Kokaram. Cepstral polynomial regression for sequential detection of impulsive waveform in video sound-track. In 1st International Conference on Semantics and Digital Media Technology, December 2006

3.17 Contribution by partner UvA

Researchers involved
Jasper Uijlings, Sennay Ghebreab, Nicu Sebe

Activities
Features extraction and testing for the Content Analysis and Segmentation Evaluation Web Portal Showcase

Achievements
Testing different features and using the existing concept detectors developed in the framework of TRECVID in order to achieve video categorization and retrieval

Events
Muscle Plenary Meeting - December 6-7-8, Paris

3.18 Contribution by partner ENSEA

Researchers involved

Sylvie Philippe-Foliguet, Michel Jordan, Frédéric Precioso, Matthieu Cord (LIP6 Partner)

Activities

Work on 3D Artwork object Indexing and Retrieving. Combining classic 3D features as input of RETIN system

Achievements

This work has been accepted in IEEE 3DTV CONFERENCE

Events

3DTV Conference 07, the true vision capture, transmission and display of 3D Video 7-9 May 2007, KICC Conference Center, Kos Island, Greece

Other

This conference is organized by the 3DTV Network of Excellence and the Informatics and Telematics Institute of Greece (ITI-CERTH).

3.19 Contribution by partner FORTH

Researchers involved

Panos Trahanias Antonis Argyros Manolis Lourakis Haris Baltzakis

Activities

FORTH has continued work on symbolic object detection and recognition, and more specifically on visual detection of objects.

FORTH has also investigated the utilization of object contours as a means for model-based 3D pose tracking. In order to extract object contours, FORTH has also continued work on methods to other sources of unutilized motion information as a means to improve results in cases that the camera is not moving. Work on the application of the developed tracking mechanisms on problems related to human/computer and human/robot interaction has also been continued.

3.20 Contribution by partner TU VIENNA-PRIP

Researchers involved

Allan Hanbury, Branislav Micusik, Lech Szumilas

Activities

Coordination of the E-team on "Choosing Features for CBIR and Automated Image Annotation". Updating of the E-team work plan at the MUSCLE general meeting in Paris.

Development of image segmentation algorithms which segment an image based on a sample of the texture to be found continues. We have worked on a new automated multi-label image segmentation approach using optimization algorithms. This technology will be included in the object recognition showcase

The investigation of image key points based on a measure of symmetry combined with a new feature describing the shape of the area around key points is under further development

Work on color interest points has been done in cooperation with the University of Amsterdam. A conference paper was submitted. This technology will be included in the object recognition showcase

Achievements
The object recognition showcase proposal was selected for funding

3.21 Contribution by partner KTH

Researchers involved
Alireza Tavakoli Targhi

Activities
Managing and updating a database of annotated images for the E-team on "Choosing Features for CBIR and Automated Image Annotation"

Achievements
The latest version of the database with 59 795 annotated images from the Corel database was provided in January 2007 at http://www.nada.kth.se/~heydarma/webpages/database.html This annotated dataset will allow evaluation of the task of recognizing animals in images. KTH provided texture features and combined texture and color features. A multiple cue feature vector was used for recognition. A database editing software had been implemented to simplify the editing task of the database much easier to enable generation of sub-databases from the original database. Partners can use this software in their own recognition and classification work. http://www.nada.kth.se/~heydarma/webpages/database_software.html

Publications
- Allan Hanbury, Alireza Tavakoli Targhi , “A Dataset of Annotated Animals. Proceedings of the Second MUSCLE / ImageCLEF Workshop on Image and Video Retrieval Evaluation (2006)

Other
The entire E-team took part in the annotation.

4 Overview activities in WP4

4.1 Contribution by partner TAU SPEECH

Researchers involved
David Burshtein, Arie Yeredor

Activities
1) Continued Algorithm development and experimentation for improved support vector machine rescoring of HMMs, with applications to speech recognition (also applies to WP6)
2) Continued experimentation and study with single-channel audio separation of speech and music through sparsifying transformations
3) Continued study of sparsifying transformations

Problems
Speech / Music Separation quality still below expectations

Events
Muscle meeting in Paris (presented part of the SVM - HMM work)

4.2 Contribution by partner TU Vienna-IFS

Researchers involved
Thomas Lidy, Rudolf Mayer

Activities
Compiled MUSCLE Inventory of Text Analysis Tools - compiled MUSCLE Inventory of Feature Extraction Tools for Audio - compiled Report on Benchmark-Based Evaluations, which describes evaluations of MUSCLE methods in international benchmarking campaigns - attended MUSCLE Plenary Meeting: gave a presentations about work in WP4 - applications and benchmarking evaluations - as well as Content-Analysis Showcase and e-Team on Semantic from Audio and Genre Classification for Music- held meeting discussing e-Team activities

Problems
Report on Benchmark-Based Evaluations delayed

Events
MUSCLE Plenary Meeting

4.3 Contribution by partner UTIA

Researchers involved
J. Grim, J. Novovicova

Activities
Text document categorization

Achievements
A subspace approach has been proposed for the problem of text categorization. By introducing a special type of Poisson mixtures including binary structural parameters we can reduce the evaluation of the Bayes formula only to a subset of informative terms/features which may be different for different classes and even for different mixture components. In this way, we can reduce the number of parameters in the conditional distributions of the documents without reducing the number of vocabulary terms.

4.4 Contribution by partner IRIT-UPS

Researchers involved

Regine Andre-Obrecht, Julien Pinquier, Khalid Daoudi, Eduardo Sanchez-Soto, Jérôme Louradour

Activities

IRIT researchers start their activities in the framework of the MUSCLE NOE at the beginning of this period. These activities are related with: - Speech, Music, Song, and Noise decomposition - Speaker verification and segmentation

4.5 Contribution by partner AUTH

Researchers involved

D. Ververidis, M. Sedaaghi, C. Kotropoulos

Activities

AUTH received MUSCLE fellow Dr. Sedaaghi who started working on feature selection using adaptive genetic algorithms in the framework of speech emotion recognition. Moreover, our research on subset feature selection using the sequential floating forward selection (SFFS) in the framework of speech emotion recognition was advanced. The criterion employed in SFFS is the prediction error of the Bayes classifier assuming that the features obey the multivariate Gaussian distribution. By employing an accurate estimate of the variance of the prediction error found by cross-validation, we have demonstrated that SFFS computational time can be reduced by 50% and the prediction error for classifying speech into emotional states for the selected subset of features varies less than the prediction error found by the usual SFFS. A theoretical analysis that models the prediction error as a binomial random variable has justified the experimental findings on Danish emotional speech (DES) and Speech Under Simulated and Actual Stress (SUSAS) databases. Our research activity is in the cross-section of Eteams 8 and 9 as well as this workpackage and workpackage 6.

Events

D. Ververidis attended the meeting at Paris in December 2006 and presented our research activities in the various Eteams AUTH is contributing to

4.6 Contribution by partner CNR-ISTI

Researchers involved

Graziano Bertini, Vincenzo Di Salvo, Thomas Lidy, Massimo Magrini, Andreas Rauber

Activities

Improvement of ARIA algorithm Winamp plug-in for transients enhancement of compressed music. Development of sw tools devoted to real-time extraction of some features of musical signals devoted to control video effects in a multimedia environment. This work is currently in progress in collaboration with private companies and supported by MIUR (Ministero Italiano Universtà e Ricerca)

Achievements

ARIA plug-in (Release 1.1)

Events
Presentation of the eteam 9 activity at MUSCLE Plenary Meeting (Paris, 6-7-8 December, 2006)

Publications
- Lunardi S., Magrini M., Tarabella L., Bertini G. - Progetto e realizzazione del controllo dell'interfaccia gestuale PalmDriver basato su microcontrollore. Internal note CNR/ISTI, B4- 13, Dec. 2006

Other
Participation of ISTI audio team in an EU Project MODEM (Contract n. I/05//B/F/PP-154059) for a "web-based environment for exchanging multi-track audio/musical signals" development

5 Overview activities in WP5

5.1 Contribution by partner UCL

Researchers involved
Fred Stentiford, Wole Oyekoya

Publications
- A thesis entitled Eye Tracking: A Perceptual Interface for Content Based Image Retrieval will shortly be publicly available

5.2 Contribution by partner TU Vienna-IFS

Researchers involved
Robert Neumayer

Activities
Wrote Book Chapter for MUSCLE Book on Multimodal Processing and Interaction - attended meeting about progress on MUSCLE Book on Multimodal Processing and Interaction

Events
MUSCLE Plenary Meeting

5.3 Contribution by partner VTT

Researchers involved
Sanni Siltanen

Activities
Participation to Muscle plenary meeting in Paris 6-8 December
Preparation of showcase proposal

Events
Muscle meeting in Paris

Other
Preparing show case -proposal

5.4 Contribution by partner ICCS-NTUA

Researchers involved

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

Activities

Audio-Visual Interaction for Speech Recognition.

Research into this field aims at improving the performance of automatic speech recognition systems in noisy environments by exploiting speech-related information extracted from video depicting the speaker's face. Audio-visual speech recognition, besides being an important research field in itself, serves as a major test-bed for methods and algorithms for cross-modal interaction potentially applicable to other multimedia integration scenarios. Part of our research on Audio-Visual speech recognition is done in collaboration with the TSI-TUC team. During the reporting period, we have launched an effort to build a real-time audio-visual automatic speech recognition demonstrator, as part of the Muscle Showcasing initiative. Participating partners are TSI-TUC (leader), ICCS-NTUA, and INRIA-TEXMEX. More specifically, experience with AV-ASR so far has been confined to research-level experimental setups: typically, videos of the speakers are shot under carefully controlled conditions, acoustic noise is artificially added, and processing is performed off-line. Towards practically deployable AV-ASR, we have been building a proof-of-concept laptop-based AV-ASR prototype which: (i) uses consumer microphone and camera to capture the speaker; (ii) performs visual/audio feature extraction, as well as speech recognition on the laptop in real-time; (iii) is robust to failures of a single modality, such as visual occlusion of the speaker's face; and (iv) automatically adapts to changing acoustic noise levels. The final system is planned to be delivered by June 2007, although earlier versions of it will be released sooner.

Achievements

Initiation of work in the "Real-Time Audio-Visual Speech Recognition" MUSCLE showcase project.

Events

Participation and presentation of research on Audio-Visual Speech Recognition at the Paris December 2006 Muscle Workshop

Publications

- A. Katsamanis, G. Papandreou, V. Pitsikalis, and P. Maragos, “Multimodal Fusion by Adaptive Compensation for Feature Uncertainty with Application to Audiovisual Speech Recognition”, in Proc. of European Signal Proc. Conf. (EUSIPCO-2006), Florence, Italy, Sep. 2006
- V. Pitsikalis, A. Katsamanis, G. Papandreou, and P. Maragos, “Adaptive Multimodal Fusion by Uncertainty Compensation”, in Proc. of Int'l Conf. Speech and Language Processing (ICSLP/INTERSPEECH-2006), Pittsburgh, PA, USA, Sep. 2006

5.5 Contribution by partner IRIT-UPS

Researchers involved

Régine André-Obrecht, Julien Pinquier, Philippe Joly, Eduardo Sanchez-Soto, Elie El-Khoury, Jeremy Philippeau, Christine Senac, Isabelle Ferrane, Zein Al-Abidin Ibrahim

Activities

IRIT researchers start their activities in the framework of the MUSCLE NOE at the beginning of this period. These activities are related with: - Person recognition based on face, costume and speech analysis, - Definition of optimal parameters for audiovisual speech recognition, - Audiovisual stream alignment on EPG - Audiovisual similarity definition and event detection.

Achievements

We finalized a non-supervised optimal stream weights computation for AV speech recognition. This work is a part of E-team 10 activities and was presented during the MUSCLE meeting in Paris (December, 6/8). Some significant studies were made on event detection in soccer programs.

Events

Participation to the Paris Meeting. Presentation of results obtained in the frameworks the E-Team10. Proposition of a showcase (ACADI N°8) on person labelling in audiovisual contents (accepted)

Publications

- Philippe Joly, "Indexation des documents audiovisuels" (Audiovisual document indexing), Habilitation Thesis, 1st of December 2006, University Paul Sabatier

5.6 Contribution by partner ICCS-NTUA

Researchers involved

P. Maragos, K. Rapantzikos, G. Evangelopoulos

Activities

Audiovisual Attention Modeling and Salient Event Detection

Although human perception appears to be automatic and unconscious there exists 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 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. The system is planned to be delivered by June 2007.

Achievements

Initiation of work in the "Movie Summarization and Skimming Demonstrator" MUSCLE showcase project

5.7 Contribution by partner ICCS-NTUA

Researchers involved

Petros Maragos

Activities

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 planned will cover the thematic areas of WP 5. It will comprise 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 WP 5 members. A rough tentative table of contents follows. Part I: State-of-the-art report(s) Merge old WP6/WP10 state of the art reports and update Part II: New research directions Possible thematic areas a.Multimodal Processing, Interaction and Understanding multimedia content i.Audio-Visual ASR ii.Feature fusion iii.Video Analysis and Integration of Asynchronous Modalities b.Audio-Visual Saliency i.Audio-Visual Scene Change and Dialogue Detection ii.Audio-Visual Attention and Salient Event Detection c.Searching multimedia content i.Annotation of multimedia databases ii.Information retrieval for video or other multimedia databases iii.Integration of Vision + Text or Audio + Text d.Interfaces to multimedia content i.Multimodal dialogue interfaces ii.Eye-tracking interfaces for information retrieval iii.Mobile interfaces.

5.8 Contribution by partner ICCS-NTUA

Researchers involved

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

Activities

Audio-Visual Speech Inversion

This research activity concerns the development of audiovisual-to-articulatory speech inversion methods. Acoustic-to-articulatory inversion is an acoustical problem that relies onto a physical model and consists of recovering geometrical data from acoustic parameters. The main difficulty is that fundamentally important aspects of the physical system of speech production, i.e. the geometry of the vocal tract, the physical characteristics of the wall, and the dependencies of the acoustic signal from the physics of speech production, cannot be

measured precisely. In addition, some of the natural data of this problem, i.e. the resonance frequencies of the vocal tract (called formants) cannot be extracted easily from speech. Alternative inversions from time-varying multiscale spectral data and/or audiovisual data are central objectives of our research because they would enable the exploitation of inversion in real applications. The acoustic data are either recorded speech or measured formant frequencies. In both cases the scientific difficulty is the infinite number of articulatory solutions. Thus, we need to deal with the incorporation of constraints that reduce the under-determination of this problem. Several aspects are investigated: the nature of information that can be incorporated (standard phonetic knowledge, audiovisual data), the most appropriate algorithmic framework and how visible constraints can be derived from video images of speaker's face. Indeed, there is strong evidence that the view of visible articulators, i.e. jaw and lips, is an important source of information for inversion. The original contributions include novel methods for automatic formant tracking, novel methods for inversion without formants' knowledge (e.g. from multiscale time-frequency information, improved articulatory speech modeling, audiovisual data), and the introduction and optimal exploitation of constraints, particularly those derived from speaker's face via appropriate video processing.

5.9 Contribution by partner FORTH

Researchers involved

Panos Trahanias, Anotnis Argyros, Haris Baltzakis

Activities

During the reporting period, FORTH was active in fields related to Multimodal Processing and Interaction. Emphasis was given to research related to recognition and interpretation of hand gestures for human/computer and human/robot interaction.

6 Overview activities in WP6

6.1 Contribution by partner CEA

Researchers involved

Moëllic Pierre-Alain, Millet Christophe, Hede Patrick

Activities

Machine Learning based on SVM techniques for the E-Team "Choosing Features for CBIR and Automated Image Annotation"

Achievements

With the features extraction works (see WP3) and SVM classification techniques, CEA reached 83.6% of correct classification for an animal detection challenge (subset of the Corel database)

6.2 Contribution by partner UU

Researchers involved

Niall Rooney

Activities

UU in conjunction with Simon Wilson wrote and submitted a showcase proposal on "A Bayesian framework for Adaptive text enhanced content based image retrieval system". The objectives of this project were: to develop an adaptive (relevance feedback-based) content based image retrieval (CBIR) system that uses both image and annotated text features, based on an existing algorithm developed at TCD by Simon Wilson that uses image features only; to develop an intuitive visual interface to allow users to select images and provide relevance feedback; to evaluate the system; to showcase the system to interested industrial parties.

6.3 Contribution by partner TU Vienna-IFS

Researchers involved

Thomas Lidy

Activities

Added a new visualization to PlaySOM software for visualizing cluster structures on a Self-Organizing Map

6.4 Contribution by partner INRIA Ariana

Researchers involved

Josiane Zerubia, Ian Jermyn, Aymen El Ghoul

Activities

Work has finished on the construction of a phase diagram for the higher-order active contour prior energy used in the network and tree crown extraction work. The calculations were finalized by Aymen El Ghoul, who began his four-month stage with INRIA Ariana in November. Mr. El Ghoul will now combine the results of the stability calculations for a circle and for a bar, in order to constrain the parameters of the model appropriately for any given application. He will also run experiments to test whether the theoretical stability calculations accord with the results of the gradient descent code used to numerically minimize the higher-order active contour energy.

6.5 Contribution by partner TUG

Researchers involved

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

Activities

1) The incremental, robust active shape model was ported from Matlab to C++. Thus, we were able to reduce the computational costs and to do extensive experiments which was not possible in Matlab.
2) By subdividing the tracker window in our object tracking framework an approximation of rotation in the integral structure is now possible with little computational effort. Therefore the deformable appearance of players during diggers, jumps and service situations can be handled.
3) Currently ensemble methods are amongst the best techniques in classification domain. These methods include particularly a very successful genre called boosting, where any member of the ensemble of classifiers are trained sequentially to compensate the shortcomings of the previously trained models, usually using the notion of sample weights. Recently, there has been a few attempts to bring the same idea of ensemble learning and

boosting to the clustering domain, but there are still theoretical difficulties in achieving well-established methods in this direction. We tried to develop a novel approach for creating partitions of data space using simple clustering algorithms in a boosting framework. A general boosting algorithm for clustering tasks has been proposed, and solutions for directly optimizing two loss functions according to this framework have been obtained. Preliminary experimental results show how the performance of relatively simple and computationally efficient base clustering algorithms could be boosted using the proposed algorithm.

6.6 Contribution by partner IBAI

Researchers involved

Petra Perner, Horst Perner

Activities

Prototype-Based Classification

Achievements

Implementation of the Method into a Program

Publications

- A Comparative Study of Catalogue-Based Classification,P. Perner, In: Th. Roth-Berghofer, M. H. Göker, H. Altay Güvenir (Eds.): Advances in Case-Based Reasoning, lncs 4106, Springer 2006, p. 301-308

6.7 Contribution by partner UCD

Researchers involved

Pádraig Cunningham, Ken Bryan

Activities

Editing book on ML for Multimedia to be published by Springer Prepared paper on Featureless Similarity for submission to ICCBR Research on Bi-clustering techniques

Achievements

Submitted paper on Featureless Similarity to ICCBR

6.8 Contribution by partner CNR-ISTI

Researchers involved

Sara Colantonio, Gabriele Pieri, Davide Moroni, Ovidio Salvetti, Massimo Chimenti, Edoardo Bozzi

Activities

A method has been developed suitable to integrate image segmentation based on active contours with cluster analysis. Preliminary tests have been done by analyzing blood cells images.

Achievements

Preparation of the article "A two-step approach for automatic microscopic image segmentation using fuzzy clustering and neural discrimination", by S. Colantonio, I.B. Gurevich, O. Salvetti, accepted for publication at PRIA Journal

Events

Seminar “Machine Learning and Image Categorization: Theory and Applications”, at DCC od RAS in Moscow

Publications

- "A two-step approach for automatic microscopic image segmentation using fuzzy clustering and neural discrimination", by S. Colantonio, I.B. Gurevich, O. Salvetti, accepted for publication at PRIA Journal

6.9 Contribution by partner ENSEA

Researchers involved

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

Activities

We propose a region-based image retrieval system, FReBIR, in which images are represented as adjacency graphs of fuzzy regions. The system is based on a module of fuzzy segmentation and a module of fuzzy region subgraph matching which retrieve images from partial queries, taking into account the image composition. Results are improved thanks to a relevance feedback which performs region classification by Support Vector Machine.

Publications

- S. Philipp-Foliguet, J. Gony, P.-H. Gosselin, FReBIR : an image retrieval system based on fuzzy region matching, submitted to CVIU, November

7 Overview activities in WP7

7.1 Contribution by partner Bilkent University

Researchers involved

A. Enis Cetin, Ozgur Ulusoy, Ugur Gudukbay, Yigithan Dedeoglu, Mehmet Turkan, R. Gokberk Cinbis, Hayati Cam

Activities

1) Multi-modal Method for Detecting Fight Among People at Unattended Places We continued our work on a surveillance system that utilizes both video and audio to detect fight among people at unattended places. In our method, first, moving objects in video are segmented from the scene background by using an adaptive background subtraction algorithm and then segmented objects are classified into groups like human and human group using a silhouette based classification method. By analyzing the motion of the human groups and at the same time detecting screams or increasing sound in audio a decision is given to detect fight.
2) Human Eye Localization we developed a human eye localization algorithm in images and video for faces with frontal pose and upright orientation with the collaboration of UPC. A given face region is filtered by a high-pass filter of a wavelet transform. In this way, edges of the region are highlighted, and a caricature-like representation is obtained. After analyzing

horizontal projections and profiles of edge regions in the high-pass filtered image, the candidate points for each eye are detected. All the candidate points are then classified using a support vector machine based classifier. Locations of each eye are estimated according to the most probable ones among the candidate points. It is experimentally observed that our eye localization method provides promising results for both image and video processing applications.
3) We continued our collaboration with ISTI-CNR about "Integration of structural and semantic models for multimedia metadata management" which is an e-team under workpackage WP2. Within this context, we will implement an MPEG-7 Feature Extractor and Query Processing Modules for Bilvideo Video Database System.

Achievements
1) We presented our results on Natural Language based Query Specification on BilVideo Video Database System in MUSCLE Scientific Meeting in Rocquencourt, Paris in December 2006
2) A summary report of E-team "Dynamic Textures and Adaptive background modelling" activities is presented in MUSCLE Scientific Meeting in Rocquencourt, Paris in December 2006

7.2 Contribution by partner MTA SZTAKI

Researchers involved
Tamas Sziranyi, Istvan Petras, Laszlo Havasi

Activities
Showcase submission: Real-time detector for unusual behaviour (e.g. falling people and fighting people)

Achievements
Organization of partners and defining the tasks Organizing the Budapest Showcase demo

Events
Paris MUSCLE meeting E-team/Showcase meeting, Barcelona

7.3 Contribution by partner SZTAKI

Researchers involved
Dmitry Chetverikov, Sandor Fazekas

Activities
A showcase proposal for dynamic texture detection in video submitted jointly with TAU-Visual.

Achievements
A pilot version of the DT detection demo created.

Events

Chetverikov and Fazekas participated in the plenary meeting in Paris where the pilot version of the DT detection demo was presented and the related showcase proposal accepted merged with the fire and smoke detection showcase proposal by Bilkent.

7.4 Contribution by partner UPC

Researchers involved

Montse Pardas, Ramon Morros, Cristian Canton

Activities

Contribution to the submission of two Showcase proposals:

- Real time detector for unusual behaviour
- Real time eyes detection system: Application to face recognition and cursor control through eye movements

7.5 Contribution by partner IBaI, CNR ISTI

Researchers involved

Petra Perner, Horst Perner, Ovido Salvetti

Activities

Advertisement of the Muscle sponsored Conference on Mass Data Analysis on Signals and Images for Medical, Biotechnological and Chemical Problems, MDA 2007 The advertisement especially involved the industry interesting in that topic. Researchers from the Muscle project are invited to submit their work for that topic and demonstrate their programs at the conference for the industry.

Achievements

There were some of the high-potential companies from the Biotechnological sector taking the MDA 2006. We hope to increase the number of companies this year by including showcase of programs from MUSCLE researchers.

Events

www.mda-signals.de

Publications

- Industrial Conference on Data Mining, ICDM 2006 Workshop Proceedings Workshop on Mass Data Analysis of Signals and Images in Medicine, Biotechnology and Chemistry MDA´2006 Workshop on Data Mining in Life Sciences Editor: Petra Perner IBaI CD-Report ISSN 1617-2671 Special Issue On Mass Data Analysis of Signals and Images in Biotechnology, Medicine, Chemistry Petra Perner (Ed.) International Journal of Signal and Imaging Systems Engineering (IJSISE) ISSN (Online): 1748-0701 - ISSN (Print): 1748-0698.

8 Resource Table

Notice that the WP-numbering refers to new WP-organisation as detailed in JPA3.

Institute	WP1a	WP1b	WP2	WP3	WP4	WP5	WP6	WP7	Total
01 - ERCIM	2.5	0	0	0	0	0	0	0	2.5
03 - UCL	0	0	0	2	0	0	0	0	2
04 - KTH	0	0	0	1	0	3	0	0	4
05 - BILKENT	0	0.2	0.5	0	0.4	0.6	0	1.9	3.6
06 - VIENNA PRIP	0	0	0.6	3	0	0	0	0	3.6
08 - UU	0	0	0	0	0	0	1	0	1
09 - CNR-ISTI	0	0.5	1.75	1.8	1.3	0	1.9	0	7.25
11 - TUG	0	0.16	0	1.5	0	0	2.4	0	4.06
12 - UPC	0	0	0	1.5	0	0	0	0.5	2
14 - UTIA	0.2	0	0	0.7	0.6	0.2	0.5	0.1	2.3
15 - UVA	0	0	2.4	0	0	0	0	0	2.4
16 - AUTH	0	0	0	0	0.614	0	0	0	0.614
18 - TU VIENNA IFS	0	0	0	0	2.8	1.2	0.55	0	4.55
19 - ACV	0	0.02	0	0.22	0	0	0	0	0.24
25 - ARMINES	0	0	0	1	0	0	0	0	1
26 - TAU-SPEECH	0	0	0	1.5	2	0	1.1	0	4.6
27 - TAU-VISUAL	0	0	0	2	0	0	0	1	3
30 - FORTH	0	0	0.22	0.21	0.11	0.22	0.22	0.66	1.64
31 - VTT	0	0	0	0	0	0.33	0	0	0.33
32 - INRIA Ariana	0	0.21	0.18	0.44	0	0.32	0.26	0	1.41
35 - UNIS	0.3	0	0	3	0	1.5	0.6	0.5	5.9
37 - ENSEA	0	0.12	0.05	0.2	0	0	0.35	0	0.72
38 - CNRS	0	0	0	0	2	0	0	0	2
39 - UPS – IRIT	0	0	0.55	0	2.2	1.75	0	0	4.5
40 - EC3	0	0	0	0	2.6	0	0	0	2.6
42 - NUID / UCD	0	0	0	0	0	0	2.2	0	2.2
Total	**0.5**	**1.21**	**7.05**	**20.47**	**15.024**	**9.32**	**11.88**	**5.66**	**71.114**