MUSCLE

Network of Excellence

Multimedia Understanding through Semantics, Computation and Learning

Project no. FP6-507752

Bi-monthly progress report # 18: Jan-Feb 2007

Due date of deliverable: 15 March 2007
Actual submission date: 20 April 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)	

Keyword List:

Contents

1 Overview activities in WP1

1.1 General scientific and administrative coordination

- Administrative and financial coordination of the network
- Preliminary preparation of the Periodic Management Report 3
- Preliminary preparation of the Periodic Activity Report 3
- Preparation of the 3rd annual EC review (3, 4 May 2007, Sophia Antipolis)
- Organisation of regular audio-conferences.
- Reimbursement of MUSCLE integration expenses
- Preparation of MUSCLE participation to the CEBIT event in March 2007 in Germany.

2 Overview activities in WP2

2.1 Contribution by TU Vienna - IFS

Researchers involved

Thomas Lidy

Activities

Evaluation of Audio Feature Extraction methods on Standard Benchmark databases

2.2 Contribution by TU VIENNA-PRIP

Researchers involved

Allan Hanbury, Branislav Micusik

Activities

Expanding the annotation of the data for the E-team on "Choosing Features for CBIR and Image Annotation". Initial work on preparing the Evaluation Showcase. The web site updates are now complete.

Publications

A. Hanbury and M. Nölle, MUSCLE CIS Coin Competition 2006, ERCIM News No. 68, page 10, January 2007. A. Hanbury and M. Nölle, Winning Algorithm for Coin Classification, IAPR Newsletter, Vol. 29, No. 1, January 2007.

2.3 Contribution by CNR-ISTI

Researchers involved

Little, Massimo Martinelli, Ovidio Salvetti, Marco Tampucc

Activities

The E-team web portal was set up for enabling collaboration and discussion. Main features are: self-editing of project pages, shared events calendar for conferences or journals of interest, shared bibliography plus public pages for publishing the e-team's goals, participants and output. A paper, titled “Integration of Structural and Semantic Models for Multimedia Metadata Management”, was written in cooperation with Bilkent University and CEA List and submitted to the CBMI conference.

The analysis of images of different typologies was started for evaluating inferencing rules and artificial neural networks as part of my MUSCLE fellowship. Improvements of the 4M architecture.

Achievements
Deliverable DN3.4, Demo of 4M.

Publications
S. Little, M. Martinelli, O. Salvetti, U. Gudukbay, O. Ulusoy, G. Grefenstette, G. de Chalendar - “Integration of Structural and Semantic Models for Multimedia Metadata Management”, submitted to CBMI conference.

2.4 Contribution by UPC

Researchers involved
Montse Pardas, Javier Ruiz, Xavier Giro

Activities
Within the e-team “Choosing features for CBIR" we have computed the following list of features for the Corel database, using the format specified in the e-team: COLOR: Layout, Structure, Mean variance SHAPE: Area, Contour, Contour-Wavelet transform, localization TEXTURE: EdgeHistogram, HomogeneousTexture

2.5 Contribution by IRIT-UPS

Researchers involved
Philippe Joly - Thomas Foures has been specially hired for this project for 4 months

Activities
A great effort concerns the development of benchmarks and the team SAMoVA has organised an evaluation campaign on video content indexing (ARGOS campaign), until February. An important study deals with distributed strategies for multimedia analysis tool integration; a platform has been developed at IRIT (project EXPRIM) and the study has given a showcase proposition TSEWP (Temporal Segmentation Evaluation Web Portal). The main activity of the SAMoVA team concerns the development of this showcase which has been merged with the "CAS" proposition to become the "CASEWP" (showcase n° 9).

2.6 Contribution by CEA

Researchers involved
Pierre-Alain MOELLIC

Activities
Preparation of the ImagEVAL workshop

Achievements
Organization of a one-day workshop after the CVIR 2007 Conference at the University of Amsterdam

2.7 Contribution by MTA SZTAKI

Researchers involved
D.Chetverikov, S. Fazekas

Activities
Performance evaluation in dynamic texture classification.

Achievements
Test results obtained and analysed.

Events
Results presented at the MUSCLE Dynamic texture workshop in Istanbul in February 2007.

Publications
Two related papers submitted, one of them already accepted (CMBI 2007).

3 Overview activities in WP3

3.1 Contribution by GET/ENST

Researchers involved
Yann Gousseau

Activities
- color image matching - geometrical image matching - mathematical modeling of natural images - indexing of satellite images - indexing of museological images

Achievements
- Automatic procedure for the querying of colour images according to their color-spatial contents. Application to various museological databases (Coll. F. Schmitt, T. Hurtut, F. Cheriet) - Resolution invariant indexing of texture-like features in satellite imaging. Application to CNES databases. - Automatic matching of sift-like features (coll. J. Delon, J. Rabin) - A contrario approach to object recognition (coll. M. Lindenbaum) - Finalisation of a work on the modeling of scaling in natural images (coll. F. Roueff)

3.2 Contribution by MTA SZTAKI, INRIA Ariana

Researchers involved
Csaba Benedek, Tamás Szirányi, Josiane Zerubia

Activities
A Three-Layer Markovian Model to Reduce Registration Errors for Object Motion Detection in Aerial Image Pairs We propose a Bayesian model for detecting the regions of object displacements in airborne image pairs taken by a moving platform. For camera motion compensation, we use a robust but coarse 2D image registration algorithm, and the main challenge is to eliminate the registration errors from the extracted change map. We introduce a three-layer Markov Random Field (MRF) model which integrates information from two different features, and ensures connected homogenous regions in the segmented images.

Achievements
The efficiency of the method has been validated through three different sets of real-world aerial images, and its behaviour versus three reference methods has been quantitatively and qualitatively evaluated.

Publications
1) Cs. Benedek, T. Szirányi: ”Color Models of Shadow Detection in Video Scenes”, Int. Conf. Computer Vision Theory and Applications (VISAPP), Barcelona, Spain, March 8-11, 2007
2) Cs. Benedek, T. Szirányi: ”Markovian Framework for Structural Change Detection with Application on Detecting Built-in Changes in Airborne Images", IASTED Int. Conf. Signal Processing, Pattern Recognition and Appl. (SPPRA 2007), Innsbruck, Austria, February 14-16, 2007

3.3 Contribution by UVA

Researchers involved
Jasper Uijlings, Sennay Ghebreab, Nicu Sebe

Activities
We investigated the use of colour information in interest point detection. Based on the Harris corner detector, we explored a way to use multi-channel images and different colour spaces were evaluated. To determine the characteristic scale of an interest point, a new colour scale selection method developed. We showed that using colour information and boosting salient colours results in improved performance in retrieval tasks. This work was performed in collaboration with TU-Vienna PRIP.

Achievements
- Publication in the Computer Vision Winter Workshop, St. Lambrecht, Austria, February 2007. - Submitted article to Int. Conference on Image Processing (ICIP 2007).

Publications
Colour Interest Points for Image Retrieval J. Stottinger, N. Sebe, T. Gevers, A. Hanbury, Computer Vision Winter Workshop, St. Lambrecht, Austria, February 2007.

3.4 Contribution by UCL

Researchers involved
Fred Stentiford, Shijie Zhang

Activities
Work Package 3.2 – Visual Saliency Work at UCL has continued to explore attention-based mechanisms. • Sponsorship from MUSCLE has enabled the world expert on Visual Attention, Professor Laurent Itti, to give a keynote presentation at the Workshop on Computational Attention and Applications in Bielefeld in March 2007. Six papers and 6 posters have been accepted for publication at the workshop - http://www.icvs2007.org/wcaa.php. • A proposal for a demonstration of attention based focusing has been submitted to the CIVR conference in July 2007.

Publications

F W M Stentiford, “Attention Based Auto Image Cropping,” Workshop on Computational Attention and Applications, ICVS07, Bielefeld, Germany, 2007. S Zhang and F W M Stentiford, “An Attention Based Method for Motion Detection and Estimation,” ICVS07, Bielefeld, Germany, 2007.

3.5 Contribution by TU VIENNA-PRIP

Researchers involved

Allan Hanbury, Branislav Micusik, Lech Szumilas, Julian Stöttinger

Activities

Coordination of the E-team on "Choosing Features for CBIR and Automated Image Annotation". Calculation of features on the dataset made available to other E-team members. We have worked on a new automated multi-label image segmentation approach using optimisation algorithms. A paper on this approach was accepted at the CVPR (Conference on Computer Vision and Pattern Recognition). 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 been done. A paper on the initial results using approach was also accepted at the CVPR conference. Work on colour interest points has been done in cooperation with the University of Amsterdam. A conference paper was presented at the Computer Vision Winter Workshop and another submitted to the International Conference on Image Processing. This technology will be included in the object recognition showcase.

Achievements

Acceptance of 2 papers at the CVPR conference - in general, this conference has an acceptance rate of around 30%.

Publications

J. Stöttinger, N. Sebe, T. Gevers and A. Hanbury, Colour Interest Points for Image Retrieval, Proceedings of the 12th Computer Vision Winter Workshop (CVWW'07), St. Lambrecht, Austria.

3.6 Contribution by TUG

Researchers involved

Martin Winter, Sandra Ober, Clemens Arth, Michael Grabner, Horst Bischof

Activities

1 Because of the fact that the visual vocabulary tree shows good results when a large number of distinctive descriptors form a large visual vocabulary and the co-occurrences perform well even on a coarse object representation with a few number of visual words we adapted the approach to a categorization approach and are currently evaluating its performance on CalTech and PASCAL- challenge object classification databases. 2. We improved our approach of treating the problem of object tracking as a matching problem of detected keypoints between successive frames. We learn discriminative classifiers for local features on-line which allows a simplification of the classification problem among the currently detected keypoints. It is possible to start tracking of objects from scratch meaning no off-line training phase is needed. An on-line boosting technique is used for learning a distance function to distinguish currently detected keypoints. Samples are collected as new frames arrive making the classifiers more and more robust over time.

This allows distinguishing learned local features from each other. The approach can be used within real-time applications since on-line updating and evaluating classifiers can be done very efficiently.

Publications
Vocabulary Tree Hypotheses and Co-Occurences; Martin Winter, Sandra Ober, Clemens Arth and Horst Bischof Proceedings of 12th Computer Vision Winter Workshop (CVWW'07), February 2007, pp.91-98

3.7 Contribution by ACV

Researchers involved
Csaba Beleznai

Activities
E-Team Meeting

Events
E-Team on Person Detection, Recognition and Tracking Meeting in Barcelona

3.8 Contribution by CNR-ISTI

Researchers involved
Umberto Barcaro, Davide Moroni, Ovidio Salvetti

Activities
Study of a novel method for shape segmentation of heart images, based on level set shape prior techniques. Preparation of the paper "A method for the automatic computation of the left ventricle ejection fraction" submitted to EUSIPCO 2007.

Achievements
Methods for left ventricle segmentation

Publications
U. Barcaro, D. Moroni, O. Salvetti; A method for the automatic computation of the left ventricle ejection fraction. Submitted to EUSIPCO 2007

3.9 Contribution by UPC

Researchers involved
Montse Pardas, Cristian Canton

Activities
Organization and participation in the two-days Workshop of e-team Person Detection, Recognition and Tracking.

Events
E-team workshop, 22nd, 23rd of January, Barcelona.

3.10 Contribution by Aristotle University of Thessaloniki

Researchers involved

Ioannis Pitas, Nikos Nikolaidis

Activities

1) Continuation of the work in facial feature detection and preparation of a journal paper.
2) Preparation of a review paper on facial feature extraction.

3.11 Contribution by IRIT-UPS

Researchers involved

Philippe Joly, Christine Sénac, Elie El Khoury (PhD student), Nourredine Besbes (Master Student)

Activities

Development of a visual-speaker detection: - Utilization and improvement of methods of the SAMoVA team based on face detection improved by the costume detection. - Utilization of a combination of the GLR and the BIC algorithms for video segmentation.

3.12 Contribution by CEA

Researchers involved

Pierre-Alain MOELLIC, Christophe MILLET, Patrick HEDE

Activities

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

Achievements

Features extraction and collecting within the HDF-5 format. The features proposed by the CEA are based on global and region-based texture, color and shape characteristics.

3.13 Contribution by MTA SZTAKI

Researchers involved

D.Chetverikov, S.Fazekas

Activities

Development of a real-time system for detection of dynamic texture in video. Development of a real-time system for video stitching. Preparing a showcase jointly with Bilkent, discussion and exchange of ideas and messages

Achievements

Pilot version of real-time system for DT detection developed and tested. Pilot version of real-time system for video stitching developed and tested.

Events

Chetverikov and Fazekas participated in the MUSCLE workshop on Dynamic Texture in Istanbul, February 2007. Two presentations given.

Publications
Joint paper with TAU-Visual accepted for presentation at the high-ranking conference SSVM 2007 (Ischia). A paper on dynamic texture recognition written and accepted for presentation in the MUSCLE session of CBMI 2007 in Bordeaux. An extended version submitted to a related special issue of a journal.

3.14 Contribution by UTIA

Researchers involved
M. Haindl, S. Mikes, G. Scarpa

Activities
Automatic detection of regions of interest in digital screening mammography

Achievements
A method for unsupervised fully automatic detection of regions of interest in digital screening mammography was developed. The unsupervised segmenter is based on a combination of several unsupervised segmentation results, each in different resolution. The mammogram tissue textures are locally represented by four monospectral MRF models evaluated for each pixel. The single-resolution segmentation part of the algorithm is based on the underlying Gaussian mixture model.
The performance of the presented method is extensively tested on the Digital Database for Screening Mammography (DDSM) from the University of South Florida as well as on the Prague segmentation benchmark.

Publications
Paper submitted to KES 2007

3.15 Contribution by GET-ENST

Researchers involved
Béatrice Pesquet-Popescu, Yann Gousseau, Christophe Tillier, Maria Trocan

Activities
Within the e-team "Visual Saliency", GET-ENST continued the exploration of adaptive and content-based multiresolution image representations, by proposing a three-step non-linear lifting scheme, involving two non-linear steps (instead of one, as we were doing previously).

Publications
Submission of a paper to IEEE ICIP conference

3.16 Contribution by TAU-SPEECH

Researchers involved
Arie Yeredor, Efrat Be'ery

Activities
Finalized the joint-diagonalization-based algorithm for image reflections separation with relative spatial shifts, submitted a paper to IEEE Transactions on Image Processing

Achievements
Separation of superimposed still-images, as well as of panned cross-fade video sequences (demo available).

Publications
Be'ery, E. and Yeredor, A., "Blind Separation of Superimposed Shifted Images Using Parameterized Joint Diagonalization" - submitted to IEEE Trans. on Image Processing.

4 Overview activities in WP4

4.1 Contribution by TU Vienna - IFS

Researchers involved
Andreas Rauber Thomas Lidy Jakob Frank

Activities
Finalized Report on Benchmark-Based Evaluations e-Team Meeting A. Rauber (TU VIENNA IFS) and Graziano Bertini (CNR) discussing progress and next steps on Dynamic optimization through music genre classification e-Team exchange visit at AIIA Lab at Aristotle University of Thessaloniki (AUTH) CASEWP Showcase Activities: Data preparation, splitting of Audio Streams, Feature Extraction, Clustering of Audio Data.

4.2 Contribution by CNRS LLACAN

Researchers involved
Fathi DEBILI, Zied BEN TAHAR, Emna SOUISSI

Activities
Morphosyntaxic analysis of Arabic. Automatic versus interactive analysis of Arabic, the proposal of metrics for the evaluation of the interactive analysis, the design and implementation of software for interactive vowelisation, the lemmatisation and POS-tagging of Arabic and lastly the evaluation.

Achievements
The development and implementation of interactive annotation algorithms (vocalisation, lemmatisation, tagging)

Publications
Submission of a paper for the 14th TALN (TALN’2007): "Analyse automatique vs analyse interactive : un cercle vertueux pour la voyellation, l’étiquetage et la lemmatisation de l’arabe". It has been accepted for an oral presentation at the TALN 2007.

4.3 Contribution by Aristotle University of Thessaloniki

Researchers involved
C. Kotropoulos

Activities

1) Work on feature/utterance elimination using self-adaptive genetic algorithms to post process the feature subset selection in emotional speech recognition. Smoothing the feature contours before applying emotional speech classification within Eteam 9.
2) Hosting Thomas Lidy of Technical University of Vienna, exchange of data and methods in order to set up joint research initiatives within Eteam 8.

4.4 Contribution by IRIT-UPS

Researchers involved

Régine Andre-Obrecht Julien Pinquier, Khalid Daoudi, Eduardo Sanchez-Soto (Post doctorant), Jerome Louradour (PhD Student), Jérôme Farinas, Hélène Lachambre (PhD Student)

Activities

- Speech, Music, Song, and Noise decomposition: a system has been developed to extract singing voice from - Speaker verification: A new "kernel model" has been theoretically studied and its application for speaker verification has been assessed (publication). This entire work is described in the PhD thesis of Jérôme Louradour. This activity is included in those of the E-Team 13 (Dynamic Kernels). - Participation to the showcase ACADI (see WP5)

Events

Jérome Louradour. « Noyaux de séquences pour la vérification du locuteur par Machines à Vecteurs de Support ». Thèse de doctorat, Université Paul Sabatier, janvier 2007

Publications

Khalid Daoudi, Jérôme Louradour. « A Novel Strategy for Speaker Verification based on SVM Classification of Pairs of Speech Sequence ». In : International Symposium on Signal Processing and its Applications (ISSPA 2007), Sharjah, United Arab Emirates, 12/02/2007-15/02/2007, IEEE, (support électronique), 2007

4.5 Contribution by CEA

Researchers involved

Gregory Grefenstette

Activities

Integration of natural language processing results into the work of the E-team "Integration of structural and semantic models for multimedia" metadata management" MUSCLE teams (BILKENT, ISTI, CEA, IBAI)

Achievements

Dissemination of this work to the wider research community has been achieved through the website http://muscle.isti.cnr.it/eteam/?q=about and the publication of a joint paper entitled "Integration of Structural and Semantic Models for Multimedia Metadata Management" at coming "Fifth International Workshop on Content-Based Multimedia Indexing" CBMI 2007 conference to take place in Bordeaux, France in June of 2007.

Events

The following members of MUSCLE participated in the organisation of the international conference on "Large Scale Semantic Access to Content (Text, Image, Video and Sound)" RIAO 2007 that will be held at Carnegie Mellon University; May 30- June 1, 2007: CEA LIST (Fluhr, Grefenstette), CWI (Pauwels), UniS (Chrsitmas), Bilkent (Duygulu), UTIA (Haindl), AUTH (Kotropoulos), UPC (Pardas), TUVienna6FS (Rauber). MUSCLE is also a co-sponsor of the event.

Publications

Techniques in building large-scale language models from the Web was presented at CICLING'2007 by the CEA in the following paper: Gregory Grefenstette. "Conquering Language: using NLP on a massive scale to build high dimensional language models from the Web" CILING'2007. Mexico, February, 2007

4.6 Contribution by TAU-SPEECH

Researchers involved

PI's: David Burshtein and Arie Yeredor

Activities

1) Continued exploration of sparsifying transformations for improved single- channel separation of audio sources. Implemented competing separation methods for comparison of performance.
2) Algorithms for support vector machine rescoring of hidden Markov models with applications to speech recognition. 3) Algorithms for efficient speaker recognition.

Publications

Submitted the following paper (partially supported by Muscle): Hagai Aronowitz and David Burshtein, "Efficient Speaker Recognition Using Approximated Cross Entropy (ACE)", Submitted to IEEE Transactions on Audio, Speech and Language Processing (Special Issue on Speaker and Language Recognition).

5 Overview activities in WP5

5.1 Contribution by TCD

Researchers involved

Rozenn Dahyot

Activities

Seminar given at Google 1 February 2007, Mountain View California, on the work on illicit material detection in video stream (work published at IET CVMP in Nov.2006 and partly funded by Muscle).

Achievements

This work is concerned with the detection of pornography in video material using both audio and visual cues.

Publications

Multimodal Periodicity Analysis for Illicit Content Detection in Videos N. Rea, G. Lacey, C. Lambe and R. Dahyot, 3rd European Conference on Visual Media Production (IET CVMP 2006), London, November 2006.

5.2 Contribution by 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.

Achievements

Progress in the "Real-Time Audio-Visual Speech Recognition" MUSCLE showcase project.

Publications

1) 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.

2) 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.

5.3 Contribution by ICCS-NTUA

Researchers involved

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

Activities

Real-time Audio-visual Automatic Speech Recognition Demonstrator Showcase

Achievements

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. During the reporting period our efforts have been directed mostly on developing the visual front-end of the prototype. The final system is planned to be delivered by June 2007, although earlier versions of it will be released sooner.

5.4 Contribution by ICCS-NTUA

Researchers involved

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

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.

5.5 Contribution by ICCS-NTUA

Researchers involved

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

Activities

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. The system is planned to be delivered by June 2007.

5.6 Contribution by ICCS-NTUA

Researchers involved

P. Maragos

Activities

Book on "Multimodal Processing and Interaction: Audio, Video, Text"
P. Maragos (ICCS-NTUA), A. Potamianos (TSI-TUC) and P. Gros (INRIA-TEXMEX), Eds.
The book will cover the thematic areas of WPs 5 and 10. 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 (former WP 6/10) members. The book proposal was submitted to Springer-Verlag and has been approved.

5.7 Contribution by ICCS-NTUA

Researchers involved

A. Katsamanis, A. Roussos, G. Papandreou, P. Maragos (ICCS),
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 modelling, audiovisual data), and the introduction and optimal exploitation of constraints, particularly those derived from speaker's face via appropriate video processing.

5.8 Contribution by FORTH

Researchers involved

Angtonis Algyros, Manolis Lourakis, Haris Baltzakis

Activities

Forth continued work FORTH on related to recognition and interpretation of hand gestures for human/computer and human/robot interaction.

5.9 Contribution by Aristotle University of Thessaloniki

Researchers involved

C. Kotropoulos, I. Pitas, N. Nikolaidis

Activities

1) Release of the AUTH Movie Database for joint research in the Movie Summarization and Skimming showcase.
2) Research in speaker clustering and face clustering in movies for dialogue detection.

5.10 Contribution by IRIT-UPS

Researchers involved

Julien Pinquier, Eduardo Sanchez-Soto, Elie El-Khoury (PhD Student),Christine Senac, Isabelle Ferrane, Zein Al-Abidin Ibrahim (PhD Student), Frederic Gianni has been specially engaged for the ACADI Showcase for 4 months -

Activities

The main activities of WP4 and WP5 concern the development of the showcase n° 8 (ACADI for Automatic Character in Audiovisual Document Indexing). It is composed of two parts: 1) Development of a system which permits to describe and structure audiovisual documents without training or corpus knowledge, and to visualize with an interface the principal interventions. We fuse three segmentation systems: face, costume and speaker detectors to obtain the best association between voices and appearing persons in an audio/video sequences. 2) Development of an interface as a tool used in a verification-aided fashion of the segmentation result. A publication of the first version of the fusion of the costum detector and the speaker detector has been submitted (accepted) to the CMBI 2007 workshop.

Achievements

For the showcase ACADI, the interface already provides the primary requirements: • Open and parse Xml file results. • Display the images of the detected characters, fetch back from the video file. • Display appearing and speaking statistics for a selected character. • Display the segmentation results for all the characters. • Merge multiple characters in one to overcome the multiple labeling of one character. • Play the video segments from the segmented sequence.

Publications

Elie El Khoury, Gaël Jaffré, Julien Pinquier, Christine Senac. Association of Audio and Video Segmentations for Automatic Person Indexing. In : International Workshop on Content-Based Multimedia Indexing (CBMI 2007), Bordeaux, France).

5.11 Contribution by TSI-TUC

Researchers involved
A. Potamianos, V. Digalakis, E. Petrakis, M. Perakakis, Th. Kanetis

Activities
Multimodal Spoken Dialogue Systems

Achievements
Evaluation results on adaptive multimodal interfaces Evaluation results on modality selection Analysis of cognitive load in MM interfaces Work in multimedia retrieval Work on fusion for AV-ASR

Publications
Journal paper under preparation on "Mode selection in MM systems"

5.12 Contribution by CWI

Researchers involved
Eric Pauwels, Romain Tavenard

Activities
Visual and audio event detection in multimodal sensor networks. Event interpretation models based on data mining of the temporal patterns in the data streams emanating from the sensors.

Publications
R.Tavenard, O. Ambekar, E. Pauwels and M. Waaijers: "Opportunistic Sensing and Learning in Sensor Networks". Proceedings. CBMI'07, Bordeaux.

5.13 Contribution by INRIA - Texmex

Researchers involved
Patrick Gros, Xavier Naturel

Activities
Xavier Naturel is finishing is thesis on TV stream structuration. This work was not reported within MUSCLE till now, but it could be interesting to use parts of this work for the movie summarization showcase. Our work was thus mainly dedicated to finding the most relevant part of this work and that of Manolis Delakis to strengthen our participation to the showcases.

5.14 Contribution by VTT

Researchers involved
Sanni Siltanen, Mika hakkrainen, Otto Korkalo

Activities
Showcase: Augmented assembly using a multimodal interface

Achievements
We did the gesture recognition for the Showcase Augmented assembly using a multimodal interface. Integrated speech recognition to showcase. And got the first version of the showcasedemo ready.

6 Overview activities in WP6

6.1 Contribution by UCD

Researchers involved
Pádraig Cunningham, Ken Bryan, Derek Greene

Activities
Editing book on ML for Multimedia to be published by Springer Prepared a paper on "An assessment of alternative strategies for constructing EMD-based kernel functions for use in an SVM for image classification" for submission to CBMI.

Achievements
The paper "An assessment of alternative strategies for constructing EMD-based kernel functions for use in an SVM for image classification" by Anton Zamolotskikh and Pádraig Cunningham has been accepted for presentation at the Fifth International Workshop on Content-Based Multimedia Indexing (CBMI 2007)

6.2 Contribution by TU Vienna - IFS

Researchers involved
Thomas Lidy Georg Pölzlbauer Rudolf Mayer Jakob Frank

Activities
Research on Classification and Kernel methods Development of PocketSOM Implementation for Mobile Devices Extended Book chapter for MUSCLE Book project on Machine Learning Techniques for Multimedia Content

Achievements
Finished Book chapter.

6.3 Contribution by ISTI CNR

Researchers involved
Emanuele Salerno, Anna Tonazzini, Ercan Kuruoglu

Activities
Research on blind separation methods for dependent sources. Within the activities of E-team 15, proposal of an invited session on computational learning methods for unsupervised segmentation (http://muscle.isti.cnr.it/CLEMUS.pdf) at the 11th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, to be held in Vietri sul Mare, Italy, from 12th to 14th September 2007. This option apppears to be preferable to the organization of an independent workshop, mainly because of the higher visibility that can be gained by the accepted contributions, which will be published in the Lecture Notes on Artificial Intelligence series.

6.4 Contribution by CNR-ISTI

Researchers involved
Sara Colantonio, Massimo Martinelli, Davide Moroni, Ovidio Salvetti

Activities
Study and development of methods for integrating inferential and computational reasoning for supporting clinical decision making

Achievements
Identification of a decision support system architecture for image-based clinical diagnosis.

6.5 Contribution by CEA

Researchers involved
Pierre-Alain MOELLIC, Christophe MILLET, Patrick HEDE

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

Achievements
Learning process tests on a larger database composed of web images of animals.

6.6 Contribution by TSI-TUC

Researchers involved
A. Potamianos, E. Iosif, I. Klasinas

Activities
Soft-clustering algorithm for automatic class induction

Achievements
We propose a soft-decision, unsupervised clustering algorithm that generates semantic classes automatically using the probability of class membership for each word, rather than deterministically assigning a word to a semantic class. Semantic classes are induced using an unsupervised, automatic procedure that uses a context-based similarity distance to measure semantic similarity between words. The proposed soft-decision algorithm is compared with various ``hard'' clustering algorithms and it is shown to improve semantic class induction performance in terms of both precision and recall for a travel reservation corpus. It is also shown that additional performance improvement is achieved by combining (auto-induced) semantic with lexical information to derive the semantic similarity distance.

Publications
Submitted to Intrespeech 2007

6.7 Contribution by UTIA

Researchers involved
J. Grim

Activities
Cluster analysis of categorical data.

Achievements
Distribution mixtures with product components have been applied repeatedly to determine clusters in multivariate data. Unfortunately, for categorical variables the mixture parameters are not uniquely identifiable and therefore the result of cluster analysis may become questionable. We give a simple proof that any non-degenerate discrete product mixture can be equivalently described by infinitely many different parameter sets. Nevertheless a unique result of cluster analysis can be guaranteed by additional constraints. We propose a heuristic method of sequential estimation of components to guarantee a unique identification of clusters by means of EM algorithm.

6.8 Contribution by UU

Researchers involved
David Patterson, Niall Rooney

Activities
We are preparing the SOPHIA system for demonstration with a number of data-sets including a crawl of WIKIPEDIA, the OHSUMED collection and the Reuters RCV1 collection.

7 Overview activities in WP7

7.1 Contribution by Bilkent University

Researchers involved
B. Ugur Toreyin, A. Enis Cetin

Activities
We study on detection and segmentation of dynamic textures in participation with our E-team: "Dynamic Textures and Adaptive Background Modeling". We improved our previous method described in `B. Ugur Toreyin and A. Enis Cetin, “Wavelet Based Detection of Moving Tree Branches and Leaves in Video”, 2006 IEEE Int. Symposium on Circuits and Systems, ISCAS 2006, Kos, Greece'.

Achievements
We presented this improved version and our IR based fire detection method at our E-team and Showcasing meeting in Istanbul, on February 12th, 2007.

Events
We organized Dynamic Textures and Adaptive Background Modelling e-team Meeting, 11-13 February 2007, Istanbul, Turkey.

7.2 Contribution by Bilkent University

Researchers involved

B. Ugur Toreyin, Yigithan Dedeoglu, A. Enis Cetin

Activities

A novel method to detect flames in infrared (IR) video is under development. Image regions containing flames appear as bright regions in IR video. In addition to ordinary motion and brightness clues, flame flicker process is also detected by using a Hidden Markov model (HMM) describing the temporal behavior. IR image frames are also analyzed spatially. Boundaries of flames are represented in wavelet domain and high frequency nature of the boundaries of fire regions is also used as a clue to model the flame flicker. All of the temporal and spatial clues extracted from the IR video are combined to reach a final decision. False alarms due to ordinary bright moving objects are greatly reduced because of the HMM based flicker modeling and wavelet domain boundary modeling.

Achievements

We presented the results of fight detection using audio and video during MUSCLE Face and Body e-team Workshop (January 22nd, 2007) at UPC in Barcelona, Spain.

7.3 Contribution by Bilkent University

Researchers involved

Ozgur Ulusoy, Ugur Gudukbay, Ediz Saykol, Ethem Fatih Can

Activities

We have been developing a scenario-based querying and retrieval model for content-based access to visual surveillance archives. In our model, a scenario can be specified as a sequence of events, and object details can be provided to enrich the queries. We have devised a visual interface as well for query specification.

Achievements

We presented our results on this research in Person Detection and Recognition, Tracking and Analysis E-team Workshop at UPC in Barcelona, Spain at January 22-23, 2007.

7.4 Contribution by Bilkent University

Researchers involved

Mehmet Turkan, A. Enis Cetin

Activities

We continued our research on Human Eye Localization.

Achievements

We presented our results on this research in Person Detection and Recognition, Tracking and Analysis E-team Workshop at UPC in Barcelona, Spain at January 22-23, 2007.

7.5 Contribution by Bilkent University

Researchers involved

Ozgur Ulusoy, Ugur Gudukbay, Onur Kucuktunc

Publications

Onur Kucuktunç, Ugur Gudukbay, Ozgur Ulusoy, "A Natural Language Based Interface for Querying a Video Database", IEEE Multimedia, Multimedia at Work, Vol. 14, No.1, pp. 83-89, 2007.

7.6 Contribution by TU Vienna - IFS

Researchers involved

Andreas Rauber, Jakob Frank, Thomas Lidy

Activities

Coordination and Preparation of MUSCLE Dissemination Activities with Showcases at CeBIT 2007 in Hannover

Events

CeBIT Hannover March 15 - 21, 2007

7.7 Contribution by Technion -MM

Researchers involved

Leonid Raskin, Ehud Rivlin, Michael Rudzsky

Activities

Research on 3D Human tracking was presented at PERSON DETECTION, RECOGNITION AND TRACKING workshop, January 07, Barcelona

Achievements

We present an approach for tracking human body parts with prelearned motion models in 3D using multiple cameras. We use an annealed particle filter to track the body parts and a Gaussian Process Dynamical Model in order to reduce the dimensionality of the problem, increase the tracker's stability and learn the motion models. We also present an improvement for the weighting function that helps to its use in occluded scenes. We compare our results to the results achieved by a regular annealed particle filter based tracker and show that our algorithm can track well even for low frame rate sequences. We also promoted a joint research with the Barcelona group from UPC - Universitat Politècnica de Catalunya on a theme Multi-camera body analysis

7.8 Contribution by ACV

Researchers involved

Csaba Beleznai, Herbert Ramoser, MTA Sztaki

Activities

Development of a tracking module for the Showcase “Real-Time Detector for Unusual Behaviour”. Meeting with MTA Sztaki to integrate tracking module into Showcase Software.

7.9 Contribution by UPC

Researchers involved
Montse Pardas, Cristian Canton, Javier Ruiz, Xavier Giro

Activities
Organization and participation in the Workshop of e-team "Person Detection, Recognition and Tracking"

Events
E-team Workshop, 22nd, 23rd of January, Barcelona

7.10 Contribution by TSI-TUC

Researchers involved
A. Potamianos, M. Perakakis, M. Maragkakis

Activities
Showcase audio-visual speech recognition demonstrator

Achievements
Front-end feature extraction for visual features Audio capture device Project management and software package version control

7.11 Contribution by TSI-TUC

Researchers involved
Potamianos, Petrakis, Drimonas,

Activities
Movie summarization showcase

Achievements
Initiated design of movie summarization interface

7.12 Contribution by MTA SZTAKI

Researchers involved
D. Chetverikov, S. Fazekas

Activities
Preparation, jointly with Bilkent, of a showcase demo on detection of fire, smoke and other dynamic textures. Preparation of a demo (video) for CeBit 2007. Contributing to a demo at the conference in Amsterdam (2007).

Achievements
A video demo for CeBit 2007 created. Real-time version of DT detector for showcase demo created.

Events
Discussions of joint showcase demo at the MUSCLE Dynamic Texture workshop in Istanbul in February 2007.

7.13 Contribution by GET-ENST

Researchers involved
Béatrice Pesquet-Popescu, Maria Trocan, Christophe Tillier

Activities
GET-ENST finished the work on fully separable wavelet and wavelet transforms as a versatile tool for dynamic texture representation. Work has started also for the showcase in collaboration with Bilkent University.

Events
Contribution to the VCIP conference, oral presentation of a paper.

Publications
M. Trocan, B. Pesquet-Popescu, Video coding with fully separable wavelet and wavelet packet transforms, VCIP 2007, SPIE Electronic Imaging, Jan 28-Feb 1st, 2007, San Jose, CA, USA

7.14 Contribution by INRIA - Texmex

Researchers involved
Guillaume Gravier, Ewa Kijak, Xavier Naturel, Patrick Gros

Activities
We began to work on three showcases: - ACADI - AV ASR - movie summarization the main events were phone conf to better specify the work to be done.

Addendum for GET activity (WP5, task2), covering period 3: March 2006 to October 2006

1) Indexing of museological images

GET : Y. Gousseau, T. Hurtut, F. Schmitt

We kept on working on the indexing and querying of images according to their color composition, in particular improved the computational efficiency of our methods. Tests have been performed on several large (up to 25000 images) artworks
databases. We also kept on developing methods to index artistic drawings, relying on multi-scale curvatures and level lines.

Publications :
Thomas Hurtut, Haroldo Dalazoana, Yann Gousseau and
Francis Schmitt, "Spatial color image retrieval without segmentation using thumbnails and the Earth Mover's Distance", CGIV 2006.

2) Natural images modeling

GET : Y. Gousseau, F. Roueff

We finalized a work on the modeling of scaling behaviors in natural images. This provides regularity indexes related to the complexity of images.
Publications :

Y. Gousseau and F. Roueff, "Modeling occlusion and scaling in natural images", SIAM Journal of Multiscale Modeling and Simulation, accepted.

C. Bordenave, Y. Gousseau and F. Roueff, "The dead leaves model : an example of a general tesselation", Adv. Appl. Probability, 38, 1, 2006, pp 31-46.

3) Statistical methods for shape recognition and image matching.

GET : J. Delon, Y. Gousseau, J. Rabin
MUSCLE collaborator : M. Lindenbaum (Technion)

A first activity has dealt with shape recognition relying on the use of topographic maps and "a contrario" methods.

A second line of research has been launched and concerns the automatic matching of local features between images using "a contrario" methods. This is investigated in two directions : single feature matching and matching of part-based model. Part of this work is performed within the e-team "statistical analysis of visual processes", in collaboration with Technion.

Publications :
F. Sur, P. Musé, F. Cao, Y. Gousseau and J.-M. Morel, "Shape recognition based on an a contrario methodology", in "Statistics and Analysis of Shapes", ed. H. Krim and A. Yezzi, Birkhauser, 2006.

F. Sur, P. Mus\'e, F. Cao, Y. Gousseau and J.-M. Morel "An a contrario decision method for shape element recognition", International Journal of Computer Vision, 69, 3, 2006, pp. 295-316.

4) Indexing of satellite images

GET : Y. Gousseau, B. Luo, H. Maitre, S. Ladjal

We kept on working on resolution independent indexing of satellite images. The work focusses on both characteristic scales and wavelet-based indexing. Test have been performed on simulated and real images provided by the French space agency.

L. Bin, J.-F. Aujol, Y. Gousseau, S. Ladjal, "Extrapolation of wavelet features for satellite images with different resolutions", proceedings of IEEE IGARSS 2006.

L. Bin, J.-F. Aujol, Y. Gousseau, S. Ladjal, H. Maitre, "Resolution independent characteristic scale in satellite images", proceedings of IEEE ICASSP, 2006.

8 Resource Table

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

Participant	WP01a	WP01b	WP02	WP03	WP04	WP05	WP06	WP07	TOTAL
01- ERCIM	2,08							0,17	2,25
02 - CWI		0,58		0,33		0,18	0,42		1,51
03 - UCL				2,00					2,00
04 - KTH					2,00	2,00			4,00
05 - BILKENT		0,20	0,30		0,25	0,40	0,20	4,50	5,85
06 - VIENNA PRIP			0,66	3,75					4,41
07 - MTA SZTAKI		0,20	1,20	0,40	0,12	0,30	1,30	0,60	4,12
08 - UU							2,30	1,80	4,10
09 - CNR-ISTI		0,06	1,90	0,30	0,25		0,10	0,10	2,71
10									
11 - TUG		0,10		1,47			2,25	0,43	4,25
12 - UPC		0,10	1,00	0,30				0,40	1,80
13 - UFR				2,00			1,00		3,00
14 - UTIA		0,20		0,80	0,60	0,10	0,40	0,20	2,30
15 - UVA				2,40					2,40
16 - AUTH				0,54		1,65			2,19
17 - CEA		0,11	0,50	0,20	0,20	0,40	0,23	0,40	2,04
18 - TU VIENNA IFS			0,15		1,75		3,60	0,50	6,00
19 - ACV				0,55				0,31	0,86
20 - TECHNION-ML							2,00		2,00
21 - TECHNION-MM				0,69	0,10	0,71		1,20	2,70
22 - IBAI		0,02	0,10	0,10	0,20				0,42
23 - ICCS		0,40	0,50	0,10	0,10	1,00		1,23	3,33
24 - TSI-TUC		0,13	0,20			1,24	0,49	1,53	3,59
25 - ARMINES				0,45	0,50				0,95
26 - TAU-SPEECH				1,00	2,50		1,00		4,50
27 - TAU-VISUAL				1,90				0,60	2,50
28 - SEIBERSDORF			1,85					0,11	1,96
29 - TCD		0,15		0,10	0,30	0,10	2,20		2,85
30 - FORTH			0,23	0,75	0,12	0,23	0,23		1,56
31 - VTT						0,56		2,00	2,56
32 - INRIA Ariana		0,26		0,41			0,26	0,24	1,17
32 - INRIA Imedia		0,50		0,72	0,05		0,30	1,00	2,57
32- INRIA Parole		0,10	0,70	0,08	0,03			0,20	1,10
32- INRIA Tex Mex						0,50		1,50	2,00
32 - INRIA Vista			0,20	0,48		0,57		0,44	1,69
33 - GET				2,40	2,00		1,31		5,71
34 - LTU			1,20						1,20
35 - UNIS				6,20		0,20	0,20		6,60
37 - ENSEA			0,10	0,25			0,35		0,70
38 - CNRS					3,00			0,20	3,20
39 - UPS – IRIT		0,10	0,35	0,40	1,20	0,85		0,05	2,95
40 - EC3		0,05			2,00		0,55		2,60
41 - UPMC		0,20					2,40		2,60
42 - NUID / UCD		0,10					2,10		2,20
Total	2,08	3,56	11,14	31,07	17,27	10,99	25,19	19,71	121,00