MUSCLE

Network of Excellence

Multimedia Understanding through Semantics, Computation and Learning

Project no. FP6-507752

Bi-Monthly Progress Report #11: Nov – Dec 2005

Due date of deliverable: 15.02.2006
Actual submission date: 20.02.2006

Start date of Project: 1 March 2004 Duration: 48 Months

Name of responsible editor(s):

- Eric Pauwels (eric.pauwels@cwi.nl), Remi Ronchaud (remi.ronchaud@cwi.nl)

Revision: 1.0

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

Keyword List:

Contents

1 Overview activities in WP 1

1.1 Contribution by ERCIM and CWI

Researchers involved: Remi Ronchaud, Eric Pauwels

Activities

- Administrative and financial coordination of the network
- Web site updating
- Launch and Stimulation of integration activities (mobility, e-teams support, conference sponsorship, joint papers)
- Organisation of the joint MUSCLE / DELOS summerschool (logisitics and scientific programme)
- Follow-up and posting of the deliverables
- Interaction with European Commission
- Preparation of the next plenary meeting in Istanbul, Turkey
- Preparation of the second Fellowship Programme (evaluations)
- Organisation of regular audio conferences
- Reimbursement of MUSCLE integration expenses (mobility support grants)

2 Overview activities in WP 2

2.1 Contribution by CWI

Researchers involved: Eric Pauwels, Margriet Brouwer

Activities Continuing development of PHP/MySQL tools for administrative and scientific reporting in WPs. Upgrade of preprint database.

3 Overview activities in WP 3

3.1 Contribution by TUVienna-PRIP

Researchers involved: Allan Hanbury,

Activities Organising MUSCLE evaluation activities in 2006. These include:

- CIS coin competition (in collaboration with ARCS)
- Some tasks in the ImageCLEF evaluation campaign

Two MUSCLE-sponsored workshops related to these activities are also being organised.

3.2 Contribution by TUVienna-PRIP

Researchers involved: Allan Hanbury, Lech Szumilas, Branislav Micusik

Activities Ground truth generation: Members of the E-team "Selecting features for CBIR and Automated Image Annotation" have been providing ground truth segmentation of animal images from the Corel database. In addition, each of the 60,000 images in the James Wang collection has been labelled as containing an animal or not.

3.3 Contribution by UTIA

Researchers involved: Michal Haindl, S.Mikes, J.Filip

Activities The Prague Texture Segmentation Datagenerator and Benchmark was modified to speed up its performance and to improve its interface. Three additional segmentation algorithms were tested on the benchmark. Among them a method developed in cooperation with INRIA Ariane. A common work on dynamic texture modelling with SZTAKI using the MUSCLE DynTex database resulted in the article submitted for ICPR 06.

3.4 Contribution by IBAI

Researchers involved: Petra Perner, Horst Perner Mike Reichle

Activities Preparation of log-Files for Data Mining Performance Analysis.

4 Overview activities in WP 4

4.1 Contribution by INRIA-Ariana

Researchers involved: Ian Jermyn, Josiane Zerubia

Activities Dissemination activities include: attendance and presentation at ICCV; a seminar at the LIAMA Institute in Beijing; establishment of the Shape Modelling E-team; attendance and presentations at the WP5/7 Focus meeting; a visit to the University of Szeged to collaborate on work related to MUSCLE, and to give a seminar.

5 Overview activities in WP 5

5.1 Contribution by TUVienna-PRIP

Researchers involved: Allan Hanbury,

Activities Collaboration with Beatriz Marcotegui of ARMINES-CMM in the framework of the E-team on "Choosing Features for CBIR and Image Annotation" on the following topics:

- Segmentation of images using the waterfall algorithm on colour-texture gradients;
- Matching of images using 2D colour histograms.

Publications MP-codes: 181

5.2 Contribution by TUVienna-PRIP

Researchers involved: Allan Hanbury, Lech Szumilas, Branislav Micusik

Activities E-team on "Choosing Features for CBIR and Automated Image Annotation"

Publications MP-codes:

5.3 Contribution by TUVienna-PRIP

Researchers involved: Allan Hanbury, Lech Szumilas, Branislav Micusik and Masters students

Activities Development of image segmentation algorithms which segment an image based on a sample of the texture to be found. This sample must be specified, for example by the user. The algorithm then attempts to mark all the regions in the image which correspond to the specified texture. This problem is an instance of the one-class classification problem, as we have information on the texture to be located, but no information on the "background" (the rest of the image). This algorithm is being further developed to be fully automatic. Furthermore, attempts to automatically locate interesting textures in an image are underway.

Publications MP-codes: 200,204,218

5.4 Contribution by AUTH

Researchers involved: Constantine Kotropoulos, Nikoletta Bassiou, Athanasios Papaioannou

Activities: Word and document clustering
Text and natural language processing

Two methods for interpolating the distanced bigram language model are examined which take into account pairs of words that appear at varying distances within a context. The language models under study yield a lower perplexity than the baseline bigram model. A word clustering algorithm based on mutual information with robust estimates of the mean vector and the covariance matrix is employed in the proposed interpolated language model. The word clusters obtained by using the aforementioned language model are proved more meaningful than the word clusters derived using the baseline bigram.

A novel method for updating probabilistic latent semantic indexing (PLSI) when new documents arrive has been developed. The proposed method adds incrementally the words of any new document in the term-document and derives the updating equations for the probability of terms given the class (i.e. latent) variables and the probability of documents given the latent variables. The performance of the proposed method is compared to that of the folding-in algorithm, which is an inexpensive but potentially inaccurate updating method. It is demonstrated that the proposed updating algorithm outperforms the folding-in method with respect to the mean squared error between the aforementioned probabilities as they are estimated by the two updating methods and the original non-adaptive PLSI algorithm. A paper on this topic has been submitted to the 4th Hellenic Conference on Artificial Intelligence.

Publications MP-codes: 81,231

5.5 Contribution by GET

Researchers involved: Beatrice Pesquet-Popescu, Christophe Tillier, Gregoire Pau

Activities We continued the work on adaptive wavelet decompositions for images. Three papers have been presented in the IEEE ICIP'05 and EURASIP EUSIPCO'05 conferences based on this work. New seminorm combinations have been tested. An interface was created to allow JPEG2000 encoding of such decompositions.

Publications MP-codes:

5.6 Contribution by ACV

Researchers involved: Herbert Ramoser, Csaba Beleznai Julia Puckmayr

Activities

- Generating annotated data for appearance based human detection
- Evaluation framework for appearance based human detection
- Annotation tool for generic ground truth creation (Matlab)
- Work on paper (with Horst Bischof from TUG): "Multiple object tracking by local PCA"

Publications MP-codes:

5.7 Contribution by MTA-SZTAKI

Researchers involved: Tamas Sziranyi, Csaba Benedek

Activities We give a new model for foreground-background-shadow separation. Our method extracts the faithful silhouettes of foreground objects even if they have partly background like colors and shadows are observable on the image. It does not need any a priori information about the shapes of the objects, it assumes only they are not point-wise. The method exploits temporal statistics to characterize the background and shadow, and spatial statistics for the foreground. A Markov Random Field model is used to enhance the accuracy of the separation. We validated our method on outdoor and indoor video sequences captured by the surveillance system of the university campus, and we also tested it on well-known benchmark videos.

Publications MP-codes: 180

5.8 Contribution by TUVienna-IFS

Researchers involved: Andreas Rauber, Thomas Lidy, Rudolf Mayer, Robert Neumayer, Rawia Awadallah

Activities: Task 3: Audio and Speech Processing

Sub-task 3.5. Events detection, segmentation and classification for audio streams

- Extended the representation and interaction model based on clusters of audio data, refacturing the demo application to allow more flexible interaction.
- Extended the visualization model to include a range of color palettes for different types of landscape.

Sub-task 3.6. High-level feature extraction for audio

Evaluated the newly developed feature representations. Specifically, compared them based on a reference data set used by the University of Athens, resulted in a first joint paper submitted and currently under review. Further steps for combining and jointly evaluating the feature sets are under discussion.

5.9 Contribution by MTA-SZTAKI

Researchers involved: Dmitry Chetverikov, Sandor Fazekas

Activities Collaboration started with TAU on developing new, computationally efficient methods for optical flow calculation which are applicable to dynamic texture recognition. Collaboration started with UTIA on synthesis and recognition of dynamic textures.

5.10 Contribution by Technion-MM

Researchers involved: Ehud Rivlin, Michael Rudzsky, Leonid Raskin

Activities: Tracking in 3D We are currently working on developing a 3D tracker. The tracker gives pose results for the target of interest. The main purpose of this tracker is to obtain 3D poses. These should support work on human action classification and analysis. We plan to extend this work to achieve part base tracking on one hand and multiple target tracking on the other hand. The part base tracking will be used to support finer action classification that commonly happens in human interaction.

5.11 Contribution by UCL

Researchers involved: Fred Stentiford, Li Chen, Oyewole Oyekoya, Adetokunbo Bamidele, Shijie Zhang

Activities Work has continued to investigate eye gaze behaviour during image search and retrieval. It has been established that under carefully controlled conditions target images can be identified more rapidly using an eyetracking interface that by using a mouse. Models of visual attention have been extended to encompass specific aspects of visual saliency such as reflective symmetry, perspective and colour constancy. New methods for determining image similarity have been combined with analysis of automatically acquired contextual metadata associated with images taken by cameraphones to substantially improve the performance of face and place recognition algorithms. Two fellowship proposals have been submitted in collaboration with CWI and INRIA. In addition an exchange of personnel is planned with INRIA to explore the application of ideas arising from saliency research to copy detection.

Publications MP-codes: 190,195-7,199,201,205-7,219-221

5.12 Contribution by UPC

Researchers involved: Montse Pardas, Cristian Canton

Activities E-team Person detection, recognition, tracking and analysis

5.13 Contribution by UPC

Researchers involved: Montse Pardas, Camilo Dorea, Ferran Marques, Veronica Vilaplana

Activities Development of image sequence segmentation, for semantic object extraction, based on color and motion features. Continue work on text detection in images and video sequences. Progress on face recognition system. Submission of 2 papers to ICIP 2006.

Publications MP-codes: 211, 213

5.14 Contribution by TUG

Researchers involved: Horst Bischof, Martina Uray, Helmut Grabner, Michael Grabner, Peter Roth

Activities

1. Construction of a common framework for object detection, tracking and recognition. It is independent of used objects and works in realtime. Progressing work comprises object identification and categorization.
2. Work on incremental LDA is continuing. Several optimization steps and the idea of retraining have been investigated. Next steps are the proof of the retraining concept and application to large scale databases.
3. A new algorithm providing learning from unlabeled data was designed. Patches for batch PCA training are provided by a tracker. The only manually interaction is the object initialisation. Expansion to realtime training and automatic initialisation is planned.

5.15 Contribution by ARMINES

Researchers involved: Beatriz Marcotegui,

Activities Segmentation of images using the waterfall algorithm on colour-texture gradients.

Publications MP-codes: 181

5.16 Contribution by ARMINES

Researchers involved: Beatriz Marcotegui,

Activities Matching of images using 2D colour histograms.

5.17 Contribution by TAU-visual

Researchers involved: Nir Sochen, Nahum Kiryati, Tammy Riklin-Raviv

Activities Developed joint post-doc proposal with INRIA-ARIANA

Publications MP-codes:

5.18 Contribution by AUTH

Researchers involved: Ioannis Pitas, Costas Cotsaces, Nikolaos Nikolaidis

Activities: Face indexing in video *Task 2: Image and video processing*
Sub-task 1. Low-level feature extraction for visual content description

A semantic video signature can be defined as a video signature based on high-level content information rather than on low-level features of the video stream. Its major advantage is that it is invariant to nearly all types of distortion. A major semantic feature of a video is the appearance of specific people in specific frames. Because of the great amount of research that has been performed on the subject of face detection and recognition, the extraction of such information is nowadays generally feasible. AUTH has developed an indexing method that uses the pre-extracted output of face detection and recognition modules to perform fast semantic indexing and retrieval of video segments. The biggest advantage of the proposed approach is that the evaluation of similarity is convolution-based, and is thus resistant to perturbations in the signature and independent of the exact boundaries of the query segment. The extraction of face information is not dealt with in this work, since ample work has been performed on the subject. The proposed method tries to solve the problems of consistency and continuity with regards to face-based indexing, to represent face information with minimal redundancy, and also to find a fast (logarithmic-time) search method. The developed method is based on reducing the information extracted regarding the existence of persons in a video to only four values:

- The time instant a person?s face starts being visible/recognizable in the video.
- The time instant a person?s face stops being visible/recognizable in the video.
- On average, how recognizable this person?s face is in this interval (which closely related to the size of the face on the video)
- Finally, the identity of the person we are dealing with (we assume the existence of a face recognition module along with the face detection module).

Each such quartet is considered as a signature element, or sigel for short. In effect, the signature of the video consists of a set of pulse series (face indicator functions), with each pulse series of the set corresponding to one person and each pulse corresponding to a sigel. In order to detect the similarity of two video segments, an integration on the pulse series obtained by evaluating the minimum (at each time instant) between two corresponding pulse series is performed. Additionally, if we want to search for the best fit between two video segments, we need to find where this integral has its maximum value when sliding one pulse series over the other. This is very similar to a convolution, except that instead of product we use the minimum operator. A significant part of the work performed so far deals with the organization of the search algorithm so that it performs with near-logarithmic efficiency with respect to video size. This is achieved by utilizing an RDBMS, and using indexes on persons and temporal intervals. It is factually impossible to process (i.e. perform face detection and recognition) the very large volume of videos (several thousand hours) which is necessary in order to test the efficiency of the proposed method, either manually or automatically. In addition, automatic processing would not allow us to control/modify the noise introduced in the semantic signatures due to face detection & recognition errors. For these reasons artificial data for the appearance and identity of persons (i.e. faces) have been generated. To achieve this, we have formulated a probabilistic model based on the analysis of the motion picture production process. We have set the parameters of this model using statistics extracted from a corpus of real video data. Preliminary results have demonstrated real-time performance in a database of 10000 artificial videos containing approximately 108 signature elements, and general robustness with respect to noise introduced from simulated face detector/recogniser errors.

The extraction of a digital signature from a video segment in order to uniquely identify it, is often a necessary prerequisite for video indexing, copyright protection, and other tasks. Semantic video signatures are those that are based on high-level content information rather than on low-level features of the video stream, their major advantage being that they are invariant to nearly all types of distortion. Since a major semantic feature of a video is the appearance of specific people in specific frames, we have developed a method that uses the pre-extracted output of face detection and recognition to perform fast semantic indexing and retrieval of video segments. We give the results of the experimental evaluation of our method on an artificial database created using a probabilistic model of the creation of video.

Publications MP-codes: 82,232

5.19 Contribution by AUTH

Researchers involved: Constantine Kotropoulos, Euthymius Ziogas, Maria Marinaki

Activities

Detection of voice disorders
Task 3: Audio and speech processing
Sub-task 3: Speech analysis

A combined scheme of linear prediction analysis is proposed for feature extraction along with linear projection methods for feature reduction followed by known pattern recognition methods on the purpose of discriminating between normal and pathological voice samples. Two different cases of speech under vocal fold pathology are examined: vocal fold paralysis and vocal fold edema. Three known classifiers are tested and compared in both cases, namely the Fisher linear discriminant, the K-nearest neighbor classifier, and the nearest mean classifier. The performance of each classifier is evaluated in terms of the probabilities of false alarm and detection or the receiver operating characteristic. The datasets used are part of a database of disordered speech developed by Massachusetts Eye and Ear Infirmary. The experimental results indicate that vocal fold paralysis and edema can easily be detected by any of the aforementioned classifiers.

A two-class pattern recognition problem has been studied, namely the automatic detection of speech disorders such as vocal fold paralysis and edema by processing the speech signal recorded from patients affected by the aforementioned pathologies as well as speakers unaffected by these pathologies. The data used were extracted from the Massachusetts Eye and Ear Infirmary database of disordered speech. The linear prediction coefficients are used as input to the pattern recognition problem. Two techniques are developed. The first technique is an optimal linear classifier design, while the second one is based on the dual-space linear discriminant analysis. Two experiments were conducted in order to assess the performance of the techniques developed namely the detection of vocal fold paralysis for male speakers and the detection of vocal fold edema for female speakers. Receiver operating characteristic curves are presented. Long-term mean feature vectors are proven very efficient in detecting the voice disorders yielding a probability of detection that may approach 100% for a probability of false alarm equal to 9.52%.

5.20 Contribution by AUTH

Researchers involved: Ioannis Pitas, Costas Cotsaces, Nikos Nikolaidis

Activities: Video shot boundary detection *Task 2: Image and video processing*
Sub-task 5: Image sequence features

A review of basic information extraction operations that can be performed on video is presented in this paper. Specifically, the review focuses on shot boundary detection and condensed video representation (also called summarization and abstraction). Shot boundary detection is the complete segmentation of a video into continuously imaged temporal video segments. Condensed video representation is the extraction of video frames or short clips that are either semantically representative of the corresponding video. Both tasks are very significant for the organization of video data into more manageable forms. An overview of the fundamental issues in each task is provided, and recent work on the subject is described and is critically reviewed.

Publications MP-codes: 237

5.21 Contribution by AUTH

Researchers involved: Constantine Kotropoulos, Dimitrios Ververidis

Activities: Emotional speech analysis and classification *Task 3: Audio and speech processing*
Subtask 5: Events detection, segmentation, and classification for audio streams

Emotional speech recognition aims to automatically classify speech units (e.g., utterances) into emotional states, such as anger, happiness, neutral, sadness and surprise. We have rated the discriminating capability of a set of features for emotional speech recognition when gender information is taken into consideration. A total of 87 features have been calculated over 500 utterances of the Danish Emotional Speech database. The Sequential Forward Selection method has been used in order to discover the 5-10 features which are able to classify the samples in the best way for each gender. The criterion used in SFS is the crossvalidated correct classification rate of a Bayes classifier where the class probability distribution functions (pdfs) are approximated via Parzen windows or modeled as Gaussians. When a Bayes classifier with Gaussian pdfs is employed, a correct classification rate of 61.1% is obtained for male subjects and a corresponding rate of 57.1% for female ones. In the same experiment, a random classification would result in a correct classification rate of 20%. When gender information is not considered a correct classification score of 50.6% is obtained.

A novel method to control the number of crossvalidation repetitions in sequential forward feature selection algorithms has been proposed. The criterion for selecting a feature is the probability of correct classification achieved by the Bayes classifier when the class feature probability density function is modeled by a single multivariate Gaussian density. Let the probability of correct classification achieved by the Bayes classifier be a random variable. We demonstrate by experiments that the probability density function of the latter random variable can be modeled by a Gaussian density. Based on this observation, a method for reducing the computational burden in Sequential forward selection algorithms is proposed. The method predicts the number of crossvalidation repetitions by employing a t-test to guarantee that a statistically significant improvement in the probability of correct classification is obtained by increasing the number of selected features. The proposed method is twice feaster than the sequential forward selection algorithm that uses a fixed number of crossvalidation repetitions and it maintains the performance of the sequential floating forward selection algorithm.

Publications MP-codes: 68,81

5.22 Contribution by AUTH

Researchers involved: Ioannis Pitas, Irene Kotsia

Activities: Dynamic facial expression recognition using support vector machines *Task 2: Image and video processing*
Sub-task 5: Image sequence features

The system consists of two subsystems, one for information extraction and one for information classification. The information extraction has been performed by a tracking system, developed based on deformable models. The system is semi-automatic, in the sense that the user has to manually place in the beginning some of the Candide wireframe model grid points to a face depicted at the first frame of the image sequence under examination. The wireframe node tracking is performed by a pyramidal variant of the well-known Kanade-Lucas-Tomasi (KLT) tracker. The loss of tracked features is handled through a model deformation procedure that increases the robustness of the tracking algorithm. The tracking system follows the facial expression evolving through time to reach its highest intensity, thus producing the grid that corresponds to it. A subset of the Candide grid points that is responsible for the formation of the movements described by the Facial Action Coding System (FACS) is chosen. The number of grid nodes finally used, is equal to 62. The facial expressions to be recognized are 6, as defined by the psychologists (anger, disgust, fear, happiness, sadness and surprise). The geometrical displacements of grid nodes, defined as the difference of each point?s coordinates between the first and the last frame of the image sequence are collected. For information classification, a six- class SVM

system is used where each class corresponds to one of the 6 basic facial expressions to be recognized. The SVM system?s kernel used is a 3rd degree polynomial function. Experiments were performed using the Cohn-Kanade database, using the leave-one-out method and the results indicate an accuracy of 97.75% with a speed of 20 frames per second for testing, therefore making it suitable for real-time applications.

Publications MP-codes: 77

5.23 Contribution by INRIA-Ariana

Researchers involved: Ian Jermyn, Josiane Zerubia

Activities Ariana's work in WP5 has focused on two areas: modelling images and modelling regions in the image domain. Together, as likelihood and prior, they enable the extraction of the regions in images corresponding to certain types of entity.

The image modelling work uses probabilistic adaptive wavelet packet models to capture texture structure, and has revealed interesting new one-point and two-point statistics for the adapted wavelet packet coefficients, statistics that are crucial for accurate texture modelling. Rather than the previous approach of using model selection, recent work uses a two-parameter family of quartic models to capture the bimodality present in many adapted subbands. The advantage of these models, aside from their greater simplicity, is that they open the way to more general models than can capture non-trivial rotation invariance and the joint statistics of pairs of bimodal subbands. A paper has been submitted to EUSIPCO 2006 on this topic.

The region modelling work is based on a new generation of active contour models, christened 'higher-order active contours', or HOACs, that enable the incorporation of non-trivial prior shape information about the region being modelled via long-range interactions between contour points, while preserving Euclidean invariance and allowing the detection of multiple instances of an entity. HOACs have been applied to the extraction of road networks from low to medium resolution satellite images.

HOACs have recently been reformulated as non-local phase field models. This reformulation offers many advantages: at the modelling level, by opening the way to probabilistic versions of the models, and hence to parameter and model estimation. A paper was published in ICCV 2005 on the phase field formulation.

Other recent work on HOACs has focused on the development of models appropriate to two problems. First, models for the extraction of road networks from very high resolution satellite images (0.5m) have begun to be developed. A joint PhD student with the LIAMA Institute in Beijing, Ting PENG (http://www-sop.inria.fr/ariana/en/lequipe.php?name=Peng) began working on this problem in September 2005, the goal being the development of a multiscale HOAC model of both the image and the road network, using the phase field reformulation of HOACs.

Moving away from the road/line network extraction problem, another joint PhD student, Peter HORVATH (http://www-sop.inria.fr/ariana/en/lequipe.php?name=Horvath), this time with the University of Szeged, has developed and analysed HOAC models of a 'gas of circles'. These models favours regions composed of a number of circles, all of approximately the same radius, and with short-range repulsion. The models are being applied to the extraction of tree crowns from aerial images. A paper has been submitted to ICPR 2006 on this topic.

The Shape Modelling e-team (http://www-sop.inria.fr/ariana/personnel/Ian.Jermyn/shapemodellingeteam/) was officially launched at the WP5/7 Focus meeting in December. A presentation on shape modelling relating to this e-team can be found at http://www-rocq.inria.fr/imedia/Muscle/WP5/WP5-FFM-docs/IanJermynArianaShapeModellingETeamPresentation.zip. The initial goals of the two founding members of the e-team, INRIA-Ariana and TAU-Visual, are centred around the incorporation of invariances, in particular affine and projective invariance, into HOAC models. More information on this and the other partners can be found at the (under construction) e-team web site (see above).

In addition, Ariana has a joint postdoc with UTIA, Giuseppe Scarpa (http://www-sop.inria.fr/ariana/en/lequipe.php?name=Scarpa). He has just begun his stay with the

Ariana project, where he will be continuing his work on the development and the evaluation of a novel unsupervised segmentation algorithm for textured images. The overall approach is based on an over-segmentation of the image, which make use of colour and contextual information, followed by a sequential region growing process, aimed at reducing the number of classes (which stops when a region gain properly defined is sufficiently large for each of the region extracted). Experiments carried out on images provided by the Prague Texture Segmentation Data-Generator and Benchmark have shown promising results.

Publications MP-codes: 245,246,250

6 Overview activities in WP 6

6.1 Contribution by INRIA-Texmex

Researchers involved: Patrick Gros, Manolis Delakis

Activities Having established a framework for audiovisual integration with Segment Models and HMMs, we explored in the past two months the possibility of fusing symbolic information with these models. In tennis videos an obvious source of symbolic information is the game structure. We used hierarchical topologies that incorporate this structure and impose constraints during decoding that the solution obtained must reflect tennis rules. Another important source of information in tennis and other sport videos is the occasionally displayed score and game statistics. A straightforward way to fuse this information is to add a descriptor that a score label is displayed to the other low-level audiovisual shot-based (or scene based, in Segment Models) features. Unfortunately, this approach cannot be efficient in videos where a lot of events are not acknowledged. Furthermore, we cannot fully exploit the symbolic information on the number of scoring events and the game evolution (number of sets and games per set) that these labels actually convey. We developed a novel search strategy which is guided by the position of these labels and using N-best-like queues we detect and penalize the paths that are not consistent with the corresponding score labels. This special Viterbi decoding thus operates in a reduced search space and guarantees that the provided solution is the most likely one and consistent with actual score and game evolution. Experimental results demonstrate a light performance improvement or degradation, depending on the models, when score labels are fused at the feature level. Instead the proposed special Viterbi decoding yields a clear performance improvement in all models.

6.2 Contribution by AUTH

Researchers involved: Constantine Kotropoulos, Margarita Kotti, Emmanouil Mpenetos, Ignacio Martinez de Lizarrondo

Activities

Speaker turn detection *Task 3: Cross-modal Integration for Multimedia Analysis and Recognition Subtask 3.1: Video Analysis and Integration of Asynschronous Time-evolving Modalities*

The activity is related to E-team 2 on Audio-Visual Understanding and in particular with the application area entitled "Dialogue detection in movies".

Unsupervised speaker change detection is a necessary step for several indexing tasks. We assume that there is no prior knowledge on the number of speakers. New features, included in the MPEG-7 Audio Prototype, are investigated such as the AudioWaveformEnvelope and the AudioSpecrtumCentroid. The model selection criterion is the Bayesian Information Criterion (BIC), which can achieve a reliable segmentation performance. A multiple pass algorithm has been developed that uses a novel dynamic thresholding and a fusion scheme so as to refine the segmentation results. The experimental

results on recordings extracted from the TIMIT database demonstrate that the performance of the proposed multiple pass algorithm is better than that of the existing approaches.

Further developments aim at speeding up the BIC-based speaker change detection by employing subset feature selection for dimensionality reduction, second-order statistical measures (such as the sphericity measures applied to the covariance matrices of the reduced features vectors, the Euclidean distance between MFCCs and the T2 Hotelling measure applied to MFFCs) before resorting to BIC.

Publications MP-codes: 235

6.3 Contribution by AUTH

Researchers involved: Constantine Kotropoulos, Ioannis Pitas, Margarita Kotti, Vassiliki Moschou, Panagiotis Antonopoulos, Spyridon Siatras

Activities: Multimodal dialogue dataset *Task 3: Cross-modal Integration for Multimedia Analysis and Recognition*
Subtask 3.1: Video Analysis and Integration of Asynschronous Time-evolving Modalities

The activity is related to E-team 2 on Audio-Visual Understanding and in particular with the application area entitled dialogue detection in movies.

A protocol to collect multimodal dataset capturing dialogues in movies is proposed. A sample multimodal dataset of 33 movie scenes extracted from 6 movies has been collected. The dataset is comprised from the audio tracks, the visual channel, and text transcriptions. Proper annotation rules employed to describe the ground truth are investigated.

6.4 Contribution by AUTH

Researchers involved: Constantine Kotropoulos, Ioannis Pitas, Margarita Kotti, Bartosz Ziolko, Vassiliki Moschou

Activities: Dialogue detection *Task 3: Cross-modal Integration for Multimedia Analysis and Recognition*
Subtask 3.1: Video Analysis and Integration of Asynschronous Time-evolving Modalities

The activity is related to E-team 2 on Audio-Visual Understanding and in particular with the application area entitled dialogue detection in movies.

We investigate two dialogue detection rules that are based on indicator functions. The first rule relies on the value of cross-correlation function at zero time lag that is compared to a threshold. The second rule is based on the cross-power in a particular frequency band that is also compared to a threshold. Experiments are carried out in order to validate the feasibility of the aforementioned dialogue detection rules by using ground-truth indicator functions determined by human observers from six different movies. Almost perfect dialogue detection is reported for every distinct threshold.

Publications MP-codes:

6.5 Contribution by AUTH

Researchers involved: Ioannis Pitas, Panagiotis Antonopoulos, Spyridon Siatras

Activities: Face clustering and mouth activity detection for dialogue detection *Task 3: Cross-modal Integration for Multimedia Analysis and Recognition*
Subtask 3.1: Video Analysis and Integration of Asynschronous Time-evolving Modalities

The activity is related to E-team 2 on Audio-Visual Understanding and in particular with the application area entitled dialogue detection in movies.

Scale invariant feature transform (SIFT) features have been used to perform face clustering on video scenes. Dialogue detection based on the successive occurrence of faces that belong in the produced clusters are investigated.

A method used to visually determine whether a person displayed in a video document is speaking using mouth intensity and edge detection information is studied.

6.6 Contribution by AUTH

Researchers involved: Constantine Kotropoulos, Marios Kyperountas, Ioannis Pitas

Activities: Audiovisual scene change detection *Task 3: Cross-modal Integration for Multimedia Analysis and Recognition*
Subtask 3.1: Video Analysis and Integration of Asynschronous Time-evolving Modalities

The activity is related to E-team 2 on Audio-Visual Understanding.

A novel audiovisual scene change detection algorithm has been developed and evaluated experimentally. An enhanced set of eigen-audioframes is extracted to create an audio signal subspace that aims at discovering the audio background changes. Visual information is used to align audio scene change indications with neighboring video shot changes and, accordingly, to reduce the false alarm rate. Moreover, video fade effects are identified and used independently in order to track scene changes. The detection methodology that processes the audio and video signals in complementary manner was tested on newscast videos provided by the TRECVID2003 video test set yielding a recall and precision rates of 80.6% and 82.1, respectively.

Publications MP-codes:

7 Overview activities in WP 7

7.1 Contribution by TCD

Researchers involved: Wilson Simon, Rozenn Dahyot

Activities Organised and attended the WP7 Focus Meeting at Rocquencourt on December 2nd. About 30 MUSCLE partners attended and there were discussions about 3 possible research collaborations (e-teams).

Publications MP-codes:

7.2 Contribution by INRIA-Ariana

Researchers involved: Ian Jermyn, Josiane Zerubia

Activities Higher-order active contours have recently been reformulated as non-local phase field models. This reformulation offers many advantages: at the algorithmic level via an immediate increase in computational efficiency, and by opening the way to further efficiency gains using multigrid, multiscale, and adaptive grid/wavelet algorithms, as well as the possibility of parallelization. A paper was published in ICCV 2005 on the phase field formulation, and Ariana hopes to have a PhD student working on this topic starting in 2006.

Publications MP-codes: 245,246,250

8 Overview activities in WP 8

8.1 Contribution by AUTH

Researchers involved: Constantine Kotropoulos, Vassiliki Moschou

Activities: Clustering by using Self-Organizing Maps

Work on clustering N-dimensional patterns that are represented as points on the (N-1)-dimensional simplex has been performed. The elements of such patterns could be the posterior class probabilities for N classes, given a feature vector derived by the Bayes classifier for example. We are interested in reducing the number of clusters to N-1, in order to redistribute the features classified into a particular class in the N-1 simplex, according to the maximum a posteriori probability principle, over the remaining N-1 classes in a optimal manner by using a self-organizing map. An application of the proposed solution to the re-assignment of emotional speech features classified as neutral into the emotional states of anger, happiness, surprise, and sadness on the Danish Emotional Speech database is presented.

The assessment of the clustering produced by two variants of the self-organizing map (SOM) that are based on order statistics, such as the marginal median SOM and the vector median SOM, has been performed. We have employed the well-known IRIS data set and we assess their performance with respect to the accuracy and the average over all neurons mean squared error between the patterns that were assigned to a neuron and the neuron's weight-vector. Both figures of merit favor the marginal median/vector median SOM against the standard SOM. Based on the aforementioned findings, the marginal median SOM and the vector median SOM have been used to re-distribute emotional speech patterns from the Danish Emotional Speech database that were originally classified as being neutral to four emotional states such as hot anger, happiness, sadness, and surprise.

Publications MP-codes: 236

8.2 Contribution by UU

Researchers involved: Niall Rooney, David Patterson

Activities We completed a study on the use of relevance feedback for textual IR system based on Contextual Document clustering, whereby we shown effectively improvements in cluster-based retrieval using few relevance judgements. We submitted a paper to Information Processing & Management on the sunject.

Publications MP-codes:

8.3 Contribution by FT

Researchers involved: Christophe Garcia,

Activities We have been carrying on our work on the development of Convolutional Neural Networks (CNN) for object detection and recognition in images: - Carried on study and development of methods for automatic building of convolutional neural network architecture, including growing and pruning techniques. - Started study of an hybrid approach: CNN and non-liner PCA for object recognition

Publications MP-codes:

8.4 Contribution by UPC

Researchers involved: Montse Pardas, Jose Luis Landabaso, Veronica Vilaplana

Activities

- Analysis of the problem of frontal face detection using Support Vector Data Description (SVDD) to characterize textural attributes of faces. The SVDD classifier relies on PCA features of face samples to obtain a decision boundary around the face data without using information of negative examples (outliers).
- Usage of Stochastic Context-Free Grammar parsers for video sequence activity classification.

Publications MP-codes: 213

8.5 Contribution by IBAI

Researchers involved: Petra Perner, Horst Perner, Silke Jänichen

Activities Refinement of conceptual clustering algorithm and writing of journal paper. Continuation of E-Team Work Unknown Pattern Recognition. Preparation of Paper on Novelty Detection and Prototype-based Classification.

Publications MP-codes:

8.6 Contribution by INRIA-Ariana

Researchers involved: Ian Jermyn, Josiane Zerubia

Activities Work has begun on parameter estimation for higher-order active contour models, and a Masters intern (http://www-sop.inria.fr/ariana/en/offres.php) has been recruited for 2006 to work on this topic.

Publications MP-codes:

8.7 Contribution by ENSEA

Researchers involved: Matthieu Cord, G. Camara sanchez, PH. Gosselin

Activities Content-based image retrieval Video cut detection Kernel methods for concept learning

Publications MP-codes: MP-252

9 Overview activities in WP9

9.1 Contribution by GET

Researchers involved: Beatrice Pesquet-Popescu, Christophe Tillier, Sebastien Brangoulo

Activities In Sept. 2005, we participated to the 74th MPEG meeting (joint with ITU/JVT) in Nice, France. In the Vidwav Ad-Hoc Group, we continued to promote the scalable video coding format based on the motion-compensated wavelet technology. We contributed to several input and output documents of this standardization body. In Dec. 2005, we participated to the 75th MPEG meeting (joint with ITU/JVT) in Bangkok, Thailand. The activity in Vidwav AHG continued, in particular with the set-up of new tests for HD sequences.

9.2 Contribution by IBAI

Researchers involved: Petra Perner, Horst Perner

Activities Transformation of logfile data into semantic understandable format

10 Overview activities in WP 10

10.1 Contribution by UCL

Researchers involved: Fred Stentiford, Wole Oyekoya, Li Chen

Activities Recent experiments have shown that visual target selection is significantly faster with an eyetracker than with a mouse. This mechanism has now been developed into an interface for image retrieval in which the user is able to search a database of images for a target image just using eye movement. A network of pre-computed inter-image similarities are used to provide successive displays of thumbnail images that are similar to the images that attracted attention in previous displays. A video (link see MUSCLE showcases) shows the fixations and saccades during a typical search. The target image is in the top left corner and the action stops as soon as the target image is retrieved.

Publications MP-codes: 195,196,197,198,199,201,205,206,219,220,221,

11 Overview activities in WP 11

11.1 Contribution by GET

Researchers involved: Beatrice Pesquet-Popescu, Maria Trocan

Activities In the framework of the E-team "3-D Texture Analysis and Detection" GET worked on using an adaptive LMS algorithm in the predict step of a lifting-based 3-D motion-compensated wavelet transform and applied this technique to scalable video coding.rnA common paper with Bilkent Univ. was submitted to the IEEE ICASSP 2006 conference.

Publications MP-codes:

11.2 Contribution by MTA-SZTAKI

Researchers involved: Tamas Sziranyi, Laszlo Havasi, Zolton Szlavik, Csaba Benedek

Activities: *Camera registration through motion Estimation of Vanishing Point in Camera-Mirror Scenes Using Video* Knowledge of the vanishing-point position is the key for the geometrical modeling of reflective surfaces or cast shadows. An automatic method is presented using motion statistics to determine correspondences, and an improved fitting function for final parameter estimation which takes into account the statistical properties of image-points. The experiments show that our approach gives robust results in the context of widely different environments especially in cases where the correspondences are corrupted with considerable amounts of noise.

Markovian Framework for Foreground-Background-Shadow Separation of Real World Video Scenes: We give a new model for foreground-background-shadow separation. Our method extracts the faithful silhouettes of foreground objects even if they have partly background like colors and shadows are observable on the image. It does not need any a priori information about the shapes of the objects,

it assumes only they are not point-wise. The method exploits temporal statistics to characterize the background and shadow, and spatial statistics for the foreground. A Markov Random Field model is used to enhance the accuracy of the separation. We validated our method on outdoor and indoor video sequences captured by the surveillance system of the university campus, and we also tested it on well-known benchmark videos.

Publications MP-codes: 187,180

11.3 Contribution by Technion-MM

Researchers involved: Ehud Rivlin, Michael Rudzsky, Leonid Raskin, Artyom Borzin

Activities

- **Human Activity Modeling for Surveillance Monitoring** We continue to develop the monitoring tool kit that allows human activity modeling using graphical user interface and video scene analysis. Currently, the proposed system integrates the Generalized Stochastic Petri Net model that will be enhanced with special handling for spatiotemporal event relations. We plan to continue extending the modeling capabilities and the training process after the system testing is completed and all strengths or weaknesses of this approach are fully analyzed. We also started working on integration of the monitoring system with recently developed tracking and classification system.
- **Human actions classification system** We are currently building a database of human actions that have been filmed with a system of four calibrated cameras. Based on that database we have started to construct a 3D pose database, which will be used in the learning phase for our pose analysis algorithm. We also have started developing a 3D pose tracker. The main purpose is to obtain 3D poses, that could be used for human action classification and analysis later on. Currently the main goal is to extend the known algorithms to be capable of tracking several persons, in order to be able to analyze human interactions.
- We have started collaboration with T. Sziranyi on Detection and Classification of Moving Objects.

11.4 Contribution by UPC

Researchers involved: Montse Pardas, Cristian Canton, Jose Luis Landabaso

Activities Development of foreground detection system oriented to real-time object tracking. Development of human action recognition system in multi-view scenarios.

Publications MP-codes: 212

11.5 Contribution by Bilkent University

Researchers involved: Yigithan Dedeoglu, Ugur Gudukbay, Serkan Genc and Enis Cetin

Activities BilVideo multimedia database system currently uses a computer vision based hand tracker to facilitate easier query creation by the motion of hand. Recently, Bilkent started working with Attila Licsar and Tamas Sziranyi (MTA SZTAKI, Hungary) to integrate their hand gesture recognition software as an input mechanism to BilVideo system. The hand gestures that can be recognized by these software systems will be used as shortcuts to various query creation actions in the BilVideo GUI. Also, the motion of the hand will replace the use of mouse for certain actions. We defined the interface between their software and BilVideo system to communicate with each other.

Publications MP-codes:

11.6 Contribution by Bilkent University

Researchers involved: B. Ugur Toreyin, Yigithan Dedeoglu, A. Enis Cetin

Activities A method for detection of tree branches and leaves in video is proposed. It is observed that the motion vectors of tree branches and leaves exhibit random motion. On the other hand regular motion of green colored objects has well-defined directions. In this study, the wavelet transform of motion vectors are computed and objects are classified according to the wavelet coefficients of motion vectors. Color information is also used to reduce the search space in a given image frame of the video. Motion trajectories of moving objects are modeled as Markovian processes and Hidden Markov Models (HMMs) are used to classify the green colored objects in the final step of the algorithm.

Publications MP-codes: 216

12 MUSCLE Publications: Papers and Preprints

For more details, consult the MUSCLE online paper archive (www.muscle-noe.org > Research > Papers).

- **MP-30:** Bouthemy, Patrick; *2D motion description and contextual motion analysis: issues and new models*
- **MP-68:** Ververidis, Dimitrios ; Kotropoulos, Constantine ; *Automatic speech classification to five emotional states based on gender information*
- **MP-69:** Marinaki, Maria; Kotropoulos, Constantine ; Pitas, Ioannis; Maglaveras, Nikolaos; *Automatic detection of vocal fold paralysis and edema*
- **MP-71:** Kovacs, Levente; Sziranyi, Tamas; *Relative Focus Map Estimation Using Blind Deconvolution*
- **MP-73:** Bassiou, Nikoletta ; Kotropoulos, Constantine ; *Interpolated distanced bigram language models for robust word clustering*
- **MP-77:** Kotsia, Irene; Pitas, Ioannis ; *Real time facial expression recognition from image sequences using Support Vector Machines*
- **MP-81:** Ververidis, Dimitrios ; Kotropoulos, Constantine ; *Sequential forward feature selection with low computational cost*
- **MP-82:** Cotsaces, Costas; Nikolaidis, Nikolaos; Pitas, Ioannis; *The use of face indicator functions for video indexing and fingerprinting*
- **MP-170:** Grim, Jiri; Somol, Petr; Haindl, Michal; Pudil, Pavel; *A Statistical Approach to Local Evaluation of a Single Texture Image*
- **MP-176:** Havasi, Laszlo; Szlavik , Zoltan ; Sziranyi, Tamas ; *Higher order symmetry for non-linear classification of human walk detection*
- **MP-177:** Kovacs, Levente; Sziranyi, Tamas; *Relative Focus Map Estimation Using Blind Deconvolution*
- **MP-179:** Benedek, Csaba; Havasi, Laszlo; Sziranyi, Tamas; Szlavik, Zoltan; *Motion-based Flexible Camera Registration*
- **MP-180:** Benedek, Csaba ; Sziranyi, Tamas; *Markovian Framework for Foreground-Background-Shadow Separation of Real World Video Scenes*
- **MP-181:** Hanbury, Allan; Marcotegui, Beatriz; *Waterfall Segmentation of Complex Scenes*
- **MP-183:** Fauqueur, Julien; Kingsbury, Nick; Anderson, Ryan; *Semantic discriminant mapping for classification and browsing of remote sensing textures and objects*
- **MP-184:** Anderson, Ryan; Kingsbury, Nick; Fauqueur, Julien ; *COARSE-LEVEL OBJECT RECOGNITION USING INTERLEVEL PRODUCTS OF COMPLEX WAVELETS*
- **MP-185:** Anderson, Ryan; Kingsbury, Nick; Fauqueur, Julien ; *Determining Multiscale Image Feature Angles from Complex Wavelet Phases*
- **MP-186:** Fauqueur, Julien; Kingsbury, Nick; Anderson, Ryan; *MULTISCALE KEYPOINT DETECTION USING THE DUAL-TREE COMPLEXWAVELET TRANSFORM*

- **MP-187:** Havasi, Laszlo ; Sziranyi, Tamas; *Estimation of Vanishing Point in Camera-Mirror Scenes Using Video*
- **MP-188:** Awadallah, Rawia; Rauber, Andreas; *Web-based Multiple Choice Question Answering for English and Arabic Questions*
- **MP-190:** Bamidele, Adetokunbo; *An attention-based model applied to colour histogram indexing*
- **MP-191:** Piella, Gemma; Pesquet-Popescu, Beatrice; Heijmans, Henk; *Gradient-Driven Update Lifting for Adaptive Wavelets*
- **MP-192:** Heijmans, Henk; Piella, Gemma; Pesquet-Popescu, Beatrice; *Adaptive Wavelets for Image Compression Using Update Lifting: Quantisation and Error Analysis*
- **MP-193:** Trocan, Maria; Tillier, Christophe; Pesquet-Popescu, Beatrice; *Joint Wavelet Packets for Groups of Frames Coding in MCTF*
- **MP-194:** Trocan, Maria; Pesquet-Popescu, Beatrice; *Scene-Cut Processing in Motion-Compensated Temporal Filtering*
- **MP-195:** Bamidele, Adetokunbo; Stentiford, Fred; *Inverted visual attention regions-of-interest for background- based search in image retrieval systems*
- **MP-196:** Stentiford, Fred; *Attention based facial symmetry detection*
- **MP-197:** Bamidele, Adetokunbo; Stentiford, Fred; *Fusing contextual metadata and visual similarity in mobile media location-based classification*
- **MP-198:** Zhang, Shijie; Stentiford, Fred; *A visual attention based method for object tracking*
- **MP-199:** Stentiford, Fred; *Attention based symmetry detection in colour images*
- **MP-200:** Kuthan, Stefan; Hanbury, Allan; *Extraction of Attributes, Nature and Context of Images*
- **MP-201:** Bamidele, Adetokunbo; Stentiford, Fred; *An attention-based similarity measure used to identify image clusters*
- **MP-202:** Piella, Gemma; Campedel, Marine; Pesquet-Popescu, Beatrice; *Adaptive Wavelets for Image Representation and Classification*
- **MP-203:** Piella, Gemma; Pau, Gregoire; Pesquet-Popescu, Beatrice; *Adaptive Lifting Schemes Combining Seminorms for Lossless Image Compression*
- **MP-204:** Seitner, Florian; Hanbury, Allan; *Fast Pedestrian Tracking based on Spatial Features and Colour*
- **MP-205:** Oyekoya, Oyewole; Stentiford, Fred; *A performance comparison of eye tracking and mouse interfaces in a target identification task*
- **MP-206:** Stentiford, Fred; *Attention based colour correction*
- **MP-207:** Davis, Marc; Smith, Michael; Stentiford, Fred; Bamidele, Adetokunbo; Canny, John; Good, Nathan; King, Simon; Janakiraman, Rajkumar; *Using Context and Similarity for Face and Location Identification*
- **MP-208:** Toreyin, Ugur; Dedeoglu, Yigithan; Cetin, Enis; *HMM Based Falling Person Detection Using Both Audio and Video*
- **MP-209:** Toreyin, Behcet Ugur; Dedeoglu, Yigithan; Gudukbay, Ugur; Cetin, A. Enis; *Computer vision based method for real-time fire and flame detection*

- **MP-210:** Dorea, Camilo; Pardas, Montse; Marques, Ferran; *A Motion-Based Binary Partition Tree Approach to Video Object Segmentation*
- **MP-211:** Dorea, Camilo; Pardas, Montse; Marques, Ferran; *Generation of lon-term color and motion coherent partitions*
- **MP-212:** Canton, Cristian; Casas, Josep; Pardas, Montse; *3D Human Action Recognition in Multiple View Scenarios*
- **MP-213:** Vilaplana, Veronica; Martinez, Claudi; Cruz, Javier; Marques, Ferran; *Face recognition using groups of images in smart room scenarios*
- **MP-214:** Vilaplana, Veronica; Marques, Ferran; *Support vector data description based on PCA features for face detection*
- **MP-215:** Toreyin, Behcet Ugur; Trocan, Maria; Pesquet-Popescu, Beatrice; Cetin, A. Enis; *LMS Based Adaptive Prediction for Scalable Video Coding*
- **MP-216:** Toreyin, Behcet Ugur; Cetin, A. Enis; *Wavelet Based Detection of Moving Tree Branches and Leaves in Video*
- **MP-217:** Grabner, Michael; Grabner, Helmut; Bischof, Horst; *Fast approximated sift*
- **MP-218:** Szumilas, Lech; Micusik, Branislav; Hanbury, Allan; *Texture Segmentation through Salient Texture Patches*
- **MP-219:** Oyekoya, Oyewole; Stentiford, Fred; *An eye tracking interface for image search*
- **MP-220:** Bamidele, Adetokunbo; Stentiford, Fred; Morphett, Jason; *An attention based approach to content based image retrieval*
- **MP-221:** Oyekoya, Oyewole; Stentiford, Fred; *Eye tracking as a new interface for image retrieval*
- **MP-222:** Louradour, Jerome; Daoudi, Khalid; Bach, Francis; *SVM Speaker Verification using an Incomplete Cholesky Decomposition*
- **MP-223:** Haasdonk, Bernard; *Transformation Knowledge in Pattern Analysis with Kernel Methods.*
- **MP-224:** Lu, Zhe-Ming; Burkhardt, Hans; *Colour Image Retrieval Based on DCT-Domain Vector Quantisation Index Histograms.*
- **MP-225:** Lu, Zhe-Ming; Li, Su-Zhi; Burkhardt, Hans; *A Content-based Image Retrieval Scheme in JPEG Compressed Domain.*
- **MP-226:** Lu, Zhe-Ming; Skibbe, Henrik; Burkhardt, Hans; *Image Retrieval Based on a Multipurpose Watermarking Scheme.*
- **MP-227:** Fehr, Janis; Ronneberger, Olaf; Kurz, Hans; Burkhardt, Hans; *Self-Learning Segmentation and Classification of Cell-Nuclei in 3D Volumetric Data using Voxel-Wise Gray Scale Invariants.*
- **MP-228:** Brunner, Gerd; Burkhardt, Hans; *Structure Features for Content-Based Image Retrieval.*
- **MP-229:** Ronneberger, Olaf; Fehr, Janis; Burkhardt, Hans; *Voxel-Wise Gray Scale Invariants for Simultaneous Segmentation and Classification.*
- **MP-230:** Reisert, Marco; Burkhardt, Hans; *Averaging Similarity Weighted Group Representations from Pose Estimation.*

- **MP-231:** Kotropoulos, Constantine; Papaioannou, Athanasios; *A novel-updating scheme for probabilistic latent semantic indexing*
- **MP-232:** Cotsaces, Costas; Nikolaidis, Nikolaos; Pitas, Ioannis; *Video indexing by face occurrence-based signatures*
- **MP-233:** Ziogas, Euthymius; Kotropoulos, Constantine ; *Detection of vocal fold paralysis and edema using linear discriminant classifiers*
- **MP-234:** Jänichen, Silke ; Perner, Petra; *Conceptual Clustering and Case Generalization of 2-dimensional Forms*
- **MP-235:** Kotti, Margarita; Benetos, Emmanouil; Kotropoulos, Constantine ; *Automatic speaker change detection with the Bayesian Information Criterion using MPEG-7 features and a fusion scheme*
- **MP-236:** Kotropoulos, Constantine; Moschou, Vassiliki; *Self-organizing maps for reducing the number of clusters by one on simplex subspaces*
- **MP-237:** Cotsaces, Costas; Nikolaidis, Nikolaos; Pitas, Ioannis; *Video shot boundary detection and condensed representation: A review*
- **MP-238:** Amiaz, Tomer; Kiryati, Nahum; *Dense Discontinuous Optical Flow via Contour-based Segmentation*
- **MP-239:** Darom, Tal; Ruggeri, Mauro; Saupe, Dietmar; Kiryati, Nahum; *Processing of Textured Surfaces Represented as Surfel Sets: Representation, Compression and Geodesic Paths*
- **MP-240:** Bar, Leah; Sochen, Nir; Kiryati, Nahum; *Semi-Blind Image Restoration Via Mumford-Shah Regularization*
- **MP-241:** Baxansky, Artemy; Kiryati, Nahum; *Calculating Geometric Properties of Three-Dimensional Objects from the Spherical Harmonic Representation*
- **MP-242:** Bar, Leah; Brook, Alexander; Sochen, Nir; Kiryati, Nahum; *Color Image Deblurring with Impulsive Noise*
- **MP-243:** Riklin-Raviv, Tammy; Sochen, Nir ; Kiryati, Nahum; *Prior-based Segmentation by Projective Registration and Level Sets*
- **MP-245:** Rochery, Marie; Jermyn, Ian; Zerubia, Josiane; *Phase field models and higher-order active contours*
- **MP-246:** Rochery, Marie; Jermyn, Ian; Zerubia, Josiane; *New Higher-order Active Contour Energies for Network Extraction*
- **MP-247:** Cossu, Roberto; Jermyn, Ian; Zerubia, Josiane; *Multimodal statistics of adaptive wavelet packet coefficients: experimental evidence and theory*
- **MP-248:** Cossu, Roberto; Jermyn, Ian; Zerubia, Josiane; *Texture discrimination using multimodal wavelet packet subbands*
- **MP-249:** Rochery, Marie; Jermyn, Ian; Zerubia, Josiane; *Gap closure in (road) networks using higher-order active contours*
- **MP-250:** Rochery, Marie; Jermyn, Ian; Zerubia, Josiane; *Higher-Order Active Contour Energies for Gap Closure*
- **MP-251:** Rochery, Marie; Jermyn, Ian; Zerubia, Josiane; *Higher Order Active Contours*

- **MP-252:** Gosselin, P.H. ; Cord, Matthieu; *Feature based approach to semi-supervised similarity learning*
- **MP-253:** Kluszczynski, , R. ; van Lieshout, M.N.M ; Schreiber, T. ; *An algorithm for binary image segmentation using polygonal Markov fields*
- **MP-254:** Ranguelova, E., Pauwels, E.J. *Saliency Detection and Matching Strategy for Photo-Identification of Humpback Whales*

13 Resource Tables

See next page.

MUSCLE - Effort Table 11 November/December 2005

(Person Month effort per Partner per Workpackage)

Participant	WP01a	WP01b	WP02	WP03	WP04	WP05	WP06	WP07	WP08	WP09	WP10	WP11	TOTAL
01- ERCIM	2,06				0,30								2,36
02 - CWI		1	0,75			1,4		0,13	1				4,28
03 - UCL						1					1		2,00
04 - KTH						1,9				0,1			2,00
05 - BILKENT		0,2	0,1	0,1	0,3	1			0,3	0,3		1,8	4,10
06 - VIENNA PRIP			0,1	0,2	0,1	2,32							2,72
07 - MTA SZTAKI				0,3		0,7			0,6			0,7	2,30
08 - UU					0,20				0,85				1,05
09 - CNR-ISTI						1		1	1	3			6,00
10 - FT		0,2				0,95			0,95				2,10
11 - TUG		0,19				1			1,3				2,49
12 - UPC					0,2	0,2			0,2			0,2	0,80
13 - UFR									1				1,00
14 - UTIA	0,1		0,1	0,5	0,4	0,5		0,3	0,4				2,30
15 - UVA							1,4						1,40
16 - AUTH						0,471	0,472						0,94
17 - CEA				0,7	0,2	0,3	0,2						1,40
18 - TU VIENNA IFS						1			0,25		0,25		1,50
19 - ACV		0,02				0,53							0,55
20 - TECHNION-ML				1					2,5	0,5			4,00
21 - TECHNION-MM					0,01	1	0,21					1	2,22
22 - IBAI		0,11	0,11	0,11	0,22				0,67	0,22			1,44
23 - ICCS		0,1		0,1	0,1	0,2	0,4				0,1		1,00
24 - TSI-TUC		0,10	0,07	0,07	0,07		0,30		0,30	0,10	1,03	0,10	2,13
25 - ARMINES		0,1		0,1		0,5							0,70
26 - TAU-SPEECH						0,6							0,60
27 - TAU-VISUAL						1,5		0,5	2				4,00
28 - SEIBERSDORF				0,2									0,20
29 - TCD		0,1		0,3		0,8		2	1,3			0,3	4,80
30 - FORTH					1	0,8	0,8					0,8	3,40
31 - VTT		0,2									0,2		0,40
32 - INRIA Ariana		0,31			0,2	0,52		0,25	0,12				1,40
32 - INRIA Imedia					0,3	0,3	0,15	0,5	0,22				1,47
32- INRIA Parole		0,2			0,2	0,5							0,90
32- INRIA Tex Mex							3						3,00
32 - INRIA Vista	0,4		0,1	0,1	0,3	1,5	0,2		0,5				3,10
33 - GET						1,84			0,89				2,73
36 - UCAM-DENG						0,10		0,10					0,20
34 - LTU						0,5							0,50
35 - UNIS					0,1								0,10
37 - ENSEA									0,3				0,30
38 - CNRS						1,5							1,50
Total	2,56	2,83	1,33	3,78	4,20	26,43	7,13	4,78	16,65	4,22	2,58	4,90	81,39