



MUSCLE

Network of Excellence

Multimedia Understanding through Semantics, Computation and Learning

Project no. FP6-507752

Bi-monthly progress report # 20: May-June 2007

Due date of deliverable: 15.07.2007

Actual submission date: 16.10.2007

Start date of project: 1 March 2004

Duration: 48 months

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

version 1.0

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

Contents

1 OVERVIEW ACTIVITIES IN WP1	3
2 OVERVIEW ACTIVITIES IN WP2	3
3 OVERVIEW ACTIVITIES IN WP3	6
4 OVERVIEW ACTIVITIES IN WP4	18
5 OVERVIEW ACTIVITIES IN WP5	22
6 OVERVIEW ACTIVITIES IN WP6	28
7 OVERVIEW ACTIVITIES IN WP7	35
8 MUSCLE EFFORT TABLE	45

1 Overview activities in WP1

1.1 General scientific and administrative coordination

- Administrative and financial coordination of the network
- Organisation of regular audio-conferences.
- E-Teams and showcase project monitoring
- Supervision and quality assessment of the deliverables
- MUSCLE Year 3 review meeting has successfully held on 3 and 4 May 2007 in Sophia Antipolis at INRIA and ERCIM facilities.
- Financial cost statements and Audit certificates were provided by UvA after repeated requests (missing in April 2007) and sent to the Commission.
- Version 2 of MUSCLE Annex I including JPA4 was sent to the project officer
- Maintenance of MUSCLE website by ERCIM
- Reimbursement of MUSCLE integration expenses (fellowship, mobility support grant)
- Preparation of MUSCLE participation to the IBC exhibition in Amsterdam in September (www.ibc.org).
- Cooperation with related European projects

2 Overview activities in WP2

2.1 Contribution by Bilkent University

Researchers involved

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

Activities & Achievements

A key goal of the MUSCLE network of excellence project is to develop tools, technologies and standards to facilitate the interoperability of multimedia content and support the exchange of such data. To this end, we formed an e-team to discuss core questions and practical issues based on the participant's individual work. Within this context, we continued our collaboration about "Integration of structural and semantic models for multimedia metadata management", with ISTI-CNR and CEA-List. We are in the process of a MPEG-7 Low level feature (color, texture, etc.) extractor for our Bilvideo Video Database System.

Publications

Suzanne Little, Massimo Martinelli, Ovidio Salvetti, Uğur Güdükbay, Özgür Ulusoy, Gael De Chalendar, Gregory Grefenstette, "Integration of Structural and Semantic Models for Multimedia Data Management", Proceedings of IEEE International Workshop on Content-Based Multimedia Indexing (CBMI'07), pp. 40-45, Bordeaux, France, June 2007

2.2 Contribution by TU VIENNA-PRIP

Researchers involved

Allan Hanbury, Branislav Micusik

Activities & Achievements

Work on the preparation of two tracks at the ImageCLEF 2007 campaign has continued. These are ImageCLEF photo, the photographic image retrieval task, and the object retrieval task. The latter task uses the PASCAL VOC 2006 data for the training set, and the IAPR-TC12 data (same as for ImageCLEF photo) for the test set. Preparation of the MUSCLE / ImageCLEF workshop to be held on the 18th of September 2007 in Budapest has continued. Two invited speakers have been secured: Thijs Westerveld, CWI (Co-organiser of the XML Multimedia Track at INEX) and Marcel Worring, University of Amsterdam (Chair of the IAPR TC12)

2.3 Contribution by CNR-ISTI

Researchers involved

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

Activities & Achievements

The work focused on further developing the 4M Infrastructure. A software package has been realized which contains all the procedures of the infrastructure and a number of new functionalities has been added. The set of image similarity queries has been broadened by enlarging the set of features that can be used for computing the similarity. The prototype for semantic annotations has been enriched by inserting the new options of adding and/or modifying the annotations to existing multimedia data. Besides the annotations, a new functionality has been included for extracting sets of significant features from a selected region. A refinement activity has also regarded the user interface, which is, in deed, under constant improvement. An investigation activity has been started for assessing different policies for uploading XML and RDF documents that detail the features extracted from images or their regions. The maintenance activity of the E-Team portal has continued. Within the fellowship program of Suzanne Little, the work concentrated on designing an hybrid approach for semantic image annotations, which is based on Machine Learning and features extraction techniques. At the end of May, the fellowship period at ISTI-CNR ended and was documented by a report on the activity carried out from September 2006 to May 2007. Within the activity aimed at developing an algorithm ontology, first steps have been performed towards the transformation of an existing thesaurus on “Image processing, analysis, recognition, and understanding” into a general purpose ontology which can be used to support a wide range of tasks, including automated image analysis, algorithmic knowledge reuse, intelligent information retrieval. A paper describing the work done has been accepted to the 2nd international conference on Semantics And digital Media Technologies (SAMT-2007), which will be held in Genoa on December 5-7 2007.

Events

Participation to the IEEE International Workshop on Content-Based Multimedia Indexing (CBMI'07), Bordeaux , June 25-27, 2007

Publications

S. Colantonio, I. Gurevich, M. Martinelli, O. Salvetti, Yu. Trusova. Thesaurus-based Ontology on Image Analysis. Accepted to the 2nd international conference on Semantics And digital Media Technologies (SAMT-2007), Genoa, December 5-7 2007

2.4 Contribution by IBaI

Researchers involved

Petra Perner Horst Perner Suzanne Little

Activities & Achievements

The work mainly concentrated on the work in the E-Team Integration of structural and semantic models for multimedia metadata management. The integration of different tools and techniques has been developed within the E-Team including for our tool ProtoClass. A report on Integrating Multimedia Metadata Management Systems has been finished. The muscle fellowship on semi-automatic semantic indexing has been further developed and a plan was made for further work at IBaI where fellow Suzanne Little is continuing her work. The results of this work so far will be published in a paper for IEEE ICDM knowledge discovery workshop.

2.5 Contribution by MTA SZTAKI

Researchers involved

S. Fazekas and D. Chetverikov (MTA SZTAKI), T.Amiaz and N.Kiryati (TAU-Visual), B.U.Toreyin, Y.Dedeoglu (Bilkent)

Activities & Achievements

Dissemination of previous results. Participation in the MUSCLE session of International Workshop on Content-Based Multimedia Indexing (CBMI), Bordeaux, 2007. A related journal publication revised and accepted. Integration of previously developed methods into a joint demo with TAU-Visual and Bilkent. Submission of the demo to the MUSCLE Showcase session of Proc. ACM International Conference on Image and Video Retrieval (CIVR), Amsterdam, 2007.

Events

Chetverikov presented results of performance evaluation of dynamic texture recognition based on optical flow and temporal periodicity at the MUSCLE session of the Fifth International Workshop on Content-Based Multimedia Indexing (CBMI), Bordeaux, 2007.

Publications

S. Fazekas and D. Chetverikov, "Analysis and performance evaluation of optical flow features for dynamic texture recognition", Special Issue of Signal Processing: Image Communication journal on Content-Based Multimedia Indexing, vol. 22/7-8, pp. 680-691, 2007.

3 Overview activities in WP3

3.1 Contribution by UCL

Researchers involved

Fred Stentiford, Shijie Zhang

Activities & Achievements

Work Package 3.2 – Visual Saliency

Work on attention based *Motion Estimation* has produced initial results which show that the method can be applied to non-rigid objects to extract the differing motion present in independently moving regions of the object. Papers are being submitted to the London Communications Symposium and the European Conference on Visual Media Production.

Work is also progressing on investigating models of *Human Visual Accommodation* and its relationship to visual attention. Experiments are being designed to compare the attention model with others by Crane, Tenengrad, an FFT and a high pass filter. An interface has been built that will allow participants to use a computer screen to select and compare blurred images as if focusing a camera. Large numbers of variously blurred images are stored in RAM to enable this control to work smoothly. Earlier work demonstrated that measures of attention are able to control the focusing of image forming devices and this work could provide a more intelligent method for future cameras. This work is being carried out in association with the Institute of Ophthalmology at UCL.

3.2 Contribution by CEA LIST

Researchers involved

Pierre-Alain Moellic Patrick Hede Christophe Millet

Activities & Achievements

CEA LIST continued research on object recognition with "animal" as a specific object class. Test have been achieved to study the performance and robustness with bag of features approaches using fast kmean clustering method for the creation of the codebook and SVM based classification. The features used were the classical SIFT descriptors. In addition, CEA LIST tests a web based approach to build automatic learning databases. The approach is based on automatic segmentation of the images provided by the Google or other search engine.

3.3 Contribution by UvA

Researchers involved

Nicu Sebe, Jasper Uijlings, Cees Snoek

Activities & Achievements

1. Adaptive discriminant analysis - Our 2D Adaptive discriminant analysis approach effectively exploits the favorable attributes of both 2D Biased Discriminant Analysis and 2D Linear Discriminant Analysis and avoids their unfavorable ones. 2D ADA can easily find an optimal discriminative subspace with adaptation to different sample distributions. It not only alleviates the problem of high dimensionality, but also enhances the classification performance in the subspace with KNN classifier. The proposed approach is applied to hand-written digit recognition and face classification. Its superior performance demonstrates that 2DADA is a promising and efficient dimension reduction approach.

2. Interactive boosting - Many content-based image retrieval applications suffer from small sample set and high dimensionality problems. Relevance feedback is often used to alleviate those problems. In this research, we propose a novel interactive boosting framework to integrate user feedback into boosting scheme and bridge the gap between high-level semantic concept and low-level image features. Our method achieves more performance improvement from the relevance feedback than AdaBoost does because human judgment is accumulated iteratively to facilitate learning process. It also has obvious advantage over the classic relevance feedback method in that the classifiers are trained to pay more attention to wrongfully predicted samples in user feedback through a reinforcement training process. An interactive boosting scheme called i.Boost is implemented and tested using Adaptive Discriminant Projection (ADP) as base classifiers, which not only combines but also enhances a set of ADP classifiers into a more powerful one. To evaluate its performance, several applications are designed on UCI benchmark data sets, Harvard, UMIST, ATT facial image data sets and COREL color image data sets. The proposed method is compared to normal AdaBoost, classic relevance feedback and the state-of-the-art projection-based classifiers. The experiment results show the superior performance of i.Boost and the interactive boosting framework.

3. Human-centered Computing - Although many researchers have worked on Human-Centered Computing (HCC) problems, and ideas around HCC have been discussed in several disciplines over the years, the field is still very young, not only in terms of research methodologies and applications, but also in terms of the definition of its scope and theoretical foundations. The definition of HCC itself remains fairly broad and what differentiates HCC research from non-HCC research is still, in many cases, an open question. The goal of our research is to investigate the challenges, the open questions, and opportunities in HCC. In our view, the multimedia community is a particularly attractive stage for discussing HCC because most multimedia applications are used directly by humans. In addition, the multidisciplinary nature of research in multimedia lends itself to an integration of concepts and techniques emanating from different assumptions and methodologies, making it an ideal ground for progress in Human-Centered Computing.

4. Emotion Recognition - Participating in the creation of a DVD on charisma in politics. Our emotion recognition software was used as one of the cues for analyzing political personalities. "Coupez le Son! Le Charisme Politique", DVD, Mona Lisa Production, Ina, April 2007.

Events

1st International workshop on Multimedia Content Analysis and Mining, WeiHai China, June 2007.

Publications

- Multimedia Content Analysis and Mining, N. Sebe, Y. Liu, Y. Zhuang, T.S. Huang, editors. Lecture Notes in Computer Science, vol. 4577, Springer, June 2007.

- Human-centered Computing: Toward a human-revolution, A. Jaimes, D. Gatica-Perez, N. Sebe, T.S. Huang, IEEE Computer, pp. 30-34, June 2007.

- Two-dimension Adaptive Discriminant Analysis Y. Lu, J. Yu, N. Sebe, Q. Tian, International Conference on Acoustics, Speech, and Signal Processing, Honolulu, April 2007.

- Integrating Relevance Feedback in Boosting for Content-Based Image Retrieval J. Yu, Y. Lu, Y. Xu, N. Sebe, Q. Tian, International Conference on Acoustics, Speech, and Signal Processing, Honolulu, April 2007.

3.4 Contribution by UPC

Researchers involved

Veronica Vilaplana, Ferran Marques

Activities & Achievements

We have developed a region-based hierarchical image representation to be used in the context of object detection. The work uses a bottom-up segmentation algorithm and, specifically, the Binary Partition Tree implementation. The different problems that arise when creating and working with a hierarchical region-based representation are analyzed; namely, (i) the creation

of the initial partition, that is, the merging and stopping criteria as well as the way to ensure a sufficient accuracy in the scene representation, (ii) the merging criteria used to build the hierarchical representation and (iii) the use of the hierarchical representation for object detection. For steps (i) and (ii), the proposed approach is assessed and compared with previous ones over a subset of the Corel database using well-established partition-based metrics. For (iii), the usefulness of the final region-based representation for object detection is exemplified in different scenarios.

Publications

Veronica Vilaplana, Ferran Marques, “Region-based hierarchical representation for object detection”, Fifth Int. Workshop on Content-based Multimedia Indexing, CBMI'07, Bordeaux, France, June 2007.

3.5 Contribution by UPC

Researchers involved

Camilo Dorea, Montse Pardas, Ferran Marques

Activities & Achievements

In this work we have developed the Trajectory Tree as a hierarchical representation for video sequences containing multiple layers of resolution and various degrees of spatiotemporal consistency. We present a technique for generating an initial set of trajectories through partition tracking. Trajectories are formed of temporally linked regions homogeneous in color and coherent in terms of affine motion models. Next, merging algorithms exploiting homogeneity criteria over multiple frames are applied towards the construction of hierarchical representations. Mechanisms for guaranteeing spatiotemporal continuity are introduced, including temporal splitting and long-term colour-based merging. The final hierarchy is constructed with the sequential application of novel long-term affine-based and translation-based motion similarity measures. The use of global motion characteristics allows the consistent segmentation and tracking of moving objects and their sub-regions, including cases of non-rigid motion and objects which stop moving for several frames.

Publications

C.C. Dórea, M. Pardàs, F. Marqués, “A hierarchical trajectory-based representation for video”, Fifth Int. Workshop on Content-based Multimedia Indexing, CBMI'07, Bordeaux, France, June 2007.

3.6 Contribution by TAU-VISUAL

Researchers involved

Ruthy Katz and Nahum Kiryati (TAU-VISUAL). Collaboration with Fred Stentiford (UCL).

Activities & Achievements

The automated puzzle solver that we develop is intended to explore and demonstrate visual attention mechanisms. The hardware is based on a PTZ (Pan-Tilt-Zoom) camera, allowing foveation processes to be implemented. A precondition to successful foveation with a PTZ camera is the ability to very accurately turn and zoom the camera to any point in the scene that is visible in the wide view of the camera. This is a significant challenge, because the PTZ camera model is very complex, and includes parameters and degrees of freedom that are not part of conventional camera models. During this period, we have made a significant progress with this problem, and are about to demonstrate it experimentally.

3.7 Contribution by TUG

Researchers involved

Martina Uray, Horst Bischof

Activities & Achievements

We did extended tests on the performance of our developed incremental LDA learning. a) Direct comparison to methods directly updating the within-class and between-class scatter matrices. b) Addition of chunks of data instead of adding one image after the other. c) Performance evaluation of different LDA methods like Fisher-LDA, GLDA or RLDA (method of our choice) in our incremental framework. d) Tests of performance decrease when reducing the number of used principal components or reducing the number of eigenvectors used to build the additional basis vectors.

Publications

Martina Uray, Danijel Skocaj, Peter Roth, Horst Bischof, Ales Leonardis "Incremental LDA Learning by Combining Reconstructive and Discriminative Approaches" British Machine Vision Conference 2007 (BMVC'07), to appear

3.8 Contribution by AUTH

Researchers involved

I. Pitas, N. Nikolaidis, I. Kotsia

Activities & Achievements

A novel method for facial expression recognition from video sequences based on fusion of texture and shape information was investigated. Regarding facial expression recognition, a subspace method based on Discriminant Non-negative Matrix Factorization (DNMF) is applied to the images, thus extracting the texture information. In order to extract the shape information, the system first extracts the deformed Candide facial grid that corresponds to the facial expression depicted in the video sequence. A Support Vector Machine (SVM) system designed on an Euclidean space, defined over a novel metric between grids, is used for the

classification of the shape information. The fusion of texture and shape information is performed using SVMs in order to detect the present facial expression. The accuracy achieved in the Cohn-Kanade database is 92.3% when recognizing the seven basic facial expressions (anger, disgust, fear, happiness, sadness, surprise, and neutral).

3.9 Contribution by TU VIENNA-PRIP

Researchers involved

Lech Szumilas, Branislav Micusik, Allan Hanbury

Activities & Achievements

We have continued work on a new automated multi-label image segmentation approach using optimisation algorithms. This work was also extended to the application of model fitting in medical images. In particular, an approach was developed that localises anatomical structures in a global manner by means of Markov Random Fields (MRF). Two papers were submitted to conferences on the latter topic.

Work on object recognition using image keypoints based on a measure of symmetry combined with a new feature describing the shape of the area around keypoints has continued. These features have also been applied to encoding salient point features in medical X-ray images in combination with the optimisation algorithms described above. Lech Szumilas has intensified cooperation with the University of Amsterdam by spending these two months doing research there.

The investigation of colour interest points in cooperation with the University of Amsterdam has continued, although this activity has mostly been shifted to the object recognition showcase in WP7.

Events

Branislav Micusik and Lech Szumilas attended and presented posters at the IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, Minnesota, USA

Publications

L. Szumilas, R. Donner, G. Langs and A. Hanbury, Local Image Structure Detection with Orientation-invariant Radial Configuration, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2007), Minneapolis, USA.

B. Micusik and T. Pajdla: Multi-label Image Segmentation via Max-sum Solver, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Minneapolis, USA, 2007.

3.10 Contribution by ISTI-CNR

Researchers involved

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

Activities & Achievements

Activity in E-team 4: A method for LV segmentation from echocardiographic image sequences has been developed. The key feature of the proposed method is the hybridization of mimetic criteria (so called because they attempt at mimicking the process followed by a human scorer) with advanced segmentation strategies. In particular, the method implements a recent variational formulation of level sets without contour re-initialization, thus circumventing the subtle problems encountered in the application of the classical level set initialization PDE. The method has been reported in a paper accepted to the Open German Russian Workshop 2007.

Events

Ercan Kuruoglu organized the Special Session “Bioinformatics applications of signal processing methods” (Ugur Sezerman and Ercan Kuruoglu Eds.), within the IEEE 15th Signal Processing and Communication Applications Conference, 11-13 June 2007, Anadolu University, Eskiehir, Turkey

Publications

U. Barcaro, D. Moroni, O. Salvetti, "Left Ventricle Segmentation in Ultrasound Sequences for the Computation of Ejection Fraction", accepted to Open German Russian Workshop, Ettlingen, Germany, 20-23 August 2007.

E.E. Kuruoglu, D. Salas, D.P. Ruiz, "Microarray gene Expression and Stable Distributions", IEEE 15th Signal Processing and Communication Applications Conference, 11-13 June 2007, Anadolu University, Eskiehir, Turkey.

K. Kayabol, E.E. Kuruoglu, B. Sankur, "Fully Bayesian Image Separation Using Markov Chain Monte Carlo", IEEE 15th Signal Processing and Communication Applications Conference, 11-13 June 2007, Anadolu University, Eskiehir, Turkey.

O. Urfalioglu, E.E. Kuruoglu, E. Cetin, "Superposed Event Detection Using Sequential Monte Carlo", IEEE 15th Signal Processing and Communication Applications Conference, 11-13 June 2007, Anadolu University, Eskiehir, Turkey.

T. Ozgen, D. Herranz and E.E. Kuruoglu, "Separation of Noisy Astrophysical Image Mixtures Using Time-Frequency Analysis Based Blind Source Separation", IEEE 15th Signal Processing and Communication Applications Conference, 11-13 June 2007, Anadolu University, Eskiehir, Turkey.

3.11 Contribution by MTA SZTAKI and INRIA Ariana

Researchers involved

Csaba Benedek, Tamas Sziranyi, Josiane Zerubia

Activities & Achievements

Common paper submitted to Journal and conference: 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.

Problems

Detecting motion on aerial camera image series

Publications

Submission of article for ICIP 2007, September 16-19.

3.12 Contribution by MTA SZTAKI and Politecnica de Catalunya

Researchers involved

István Petrás Montse Pardàs Levente Kovács László Havasi Tamás Szirányi Cristian Canton

Activities & Achievements

Human model and motion based unusual event detection In order to achieve a simple motion representation, a Motion History Image (MHI) and Motion Energy Image (MEI) are applied. This representation has been recently used for monocular gait recognition tasks and activity modeling. We have extended this formulation to represent view-independent 3D motion. A simple ellipsoid body model was fit to the incoming 3D data to capture in which body part the gesture occurs thus increasing the recognition ratio of the overall system and generating a more informative classification output. In this module, we use the same approach for monocular sequences. The system is based on three sub-modules.

Problems

Detection of moving body

Events

Participation to CIVR 07 and visit to Budapest by UPC partner

Publications

CIVR 07 in July 2007, Amsterdam

3.13 Contribution by LTU Technologies

Researchers involved

S. Gilles, P. Chiquet

Activities & Achievements

We have developed a new video copy detection technique with 2 major goals: - have a scalable approach that scales to thousands of hours - have a light-, contrast-, color-tolerant approach. A video is indexed by computing a temporal signature. This temporal signature itself is obtained by indexing selected frames only, exploiting inter-frame redundancy. Each frame is indexed using a set of integral features, whose binary outputs are then combined using a linear function to produce a N-bit word. The video signature is then the consecutive series of frame signatures. The matching problem is addressed as a partial pattern matching problem: given a string of T characters with potentially E errors, find the best-matching substring among a set of strings of varying lengths. To make the problem tractable, a probabilistic framework updates at each new indexed frame a set of probabilities. Each probability representing a potential time-constrained association between the unknown and a candidate match. Depending on a wise choice of features, the system shows rapid convergence towards the correct solution if a match exists in the signature database, or exhibits a flat set of probabilities when no match exists in the database. Tolerance is very strong when it comes to photometric artefacts (local and global contrast, color and lighting changes), to video re-encoding, to video resizing, to frame rate changing. The system is also able to retrieve videos even in the presence of subtitles, small crops & inclusion.

Problems

We took part in the white test of the video benchmark organized for CIVR2007 and the system exhibit excellent quality and response times. Participation to the actual benchmark in Amsterdam in July could however not be made possible, due to tight development times. LTU plans on participating to the next international benchmark on video copy detection.

3.14 Contribution by INRIA Ariana

Researchers involved

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

Activities & Achievements

Ting Peng, joint PhD student of INRIA Ariana and the LIAMA Institute in Beijing, has continued to work on models for road network segmentation from very high resolution (0.5m) satellite images. As described in the previous report, having developed a successful model for the extraction of main roads, Ms Peng has been working on extending these models to narrower roads. Ms Peng has developed three models for the latter. Two involve asymmetrizing the interaction in the higher-order energy term to allow it to vary with the relative orientation of the contour normal vectors. The third uses a hitherto unrecognized linear term that takes into account the relative orientation of the vector between the two interacting points and the normal vectors. The idea of all three models is to enforce road edge straightness over a larger range than the road width, although they do this in different ways. Ms Peng now has the first experimental results based on one of the nonlinear models, and the results are remarkable. The edges of the roads are more accurately extracted than before and narrower roads are more easily handled. Junctions are crisper, with less smoothing of right angles than was the case with the previous linear model. Further experimental testing is required to examine the generalization ability of the model, and comparative experiments must be performed for the other two models, in particular the new linear model, since this is less intensive computationally and is therefore preferable if the quality of the results is comparable. Aymen El Ghoul, student at Sup'Com, Tunis, is in the middle of a second internship with INRIA Ariana as a prelude to taking up a PhD position in the group later in the year. He is working on extending the higher-order active contour model of road networks, and its phase field formulation, to the cases of grid-like road networks and river networks, with the aim of extracting these entities from high resolution satellite and aerial images. Having performed a bibliographic study of possible modelling tools, and an assessment of the state of the art for these problems, he submitted an internal report in June 2007. He has also performed a stability analysis for a long straight bar analogous to that performed by Mr. Horvath for a circle, and has tested it against the results of gradient descent with success. This analysis will lead to constraints on the model parameters for network models. He has submitted a paper to RFIA 2008 on this topic, which is in the process of being transformed into an INRIA Research Report and eventually a journal publication. In view of the ability of the model developed by Ms Peng to extract clean right angles, as described above, it seems that this might be a good starting point for a model to extract grid-like road networks from urban areas. Mr. El Ghoul is currently performing a stability analysis for this new energy similar to those performed for the previous linear energy, with the aim of better understanding its behaviour. Peter Horvath, joint PhD student of INRIA Ariana and the University of Szeged, has been writing up his thesis, and will submit it to the University of Szeged in August, and to the University of Nice Sophia Antipolis in September. Avik Bhattacharya, joint PhD student of INRIA Ariana and ENST, has now finished extensive tests to compare his features and method with other existing feature sets, with good results. He has started to write up his thesis, in order to submit it to the University of Nice Sophia Antipolis in August or September.

In addition, scientific coordination and advising for all interns, Masters students, PhD students, and post-docs working in or with INRIA Ariana on MUSCLE topics was achieved. Coordination of MUSCLE e-team 'Shape Modelling' including budget and funding applications was performed. All publications related to INRIA Ariana's work within the Muscle project may be found at <http://www-sop.inria.fr/ariana/en/publications.php> by searching for the appropriate researcher.

Publications

A paper has been accepted at the conference BMVC 2007 (Peng).

3.15 Contribution by ENSEA

Researchers involved

D. Gorisse, Prof. S. Philipp-Foliguet, M. Jordan, Dr. F. Precioso (ENSEA), Prof. M. Cord (UPMC)

Activities & Achievements

We have developed an algorithm of segmentation of the 3D object surface: at first, we calculate the values of the local curvatures at each point of the mesh, then these values are introduced into an algorithm of watersheds to perform the segmentation of the 3D surface. Features are then computed for every region ; these features are the same as those computed for complete models: cord histograms, extended Gaussian images, etc...[Gorisse 2007]. They characterize the regions, and will allow partial queries (one or several regions of the model) in a database of 3D objects in order to retrieve objects 3D having similar parts, or objects from a single fragment.

Publications

[Gorisse 2007] D. Gorisse, M. Cord, M. Jordan, S. Philipp-Foliguet, F. Precioso, 3D Content-Based Retrieval in Artwork Databases, Proceedings of the 3DTV-Conference, 7-9 May 2007, Kos Island, Greece.

3.16 Contribution by GET/ENST

Researchers involved

Yann Gousseau

Activities & Achievements

1) Matching of local features (Coll. J. delon and J. Rabin) - We kept on investigating the use of "a contrario" method for local features matching. This period work was concerned with clustering issues. A conference paper has been accepted (Gretsi). A journal paper is in preparation. 2) Indexing of satellite images (Coll. J.-F. Aujol, L. Bin and H. Maitre) We kept on working on the indexing of satellite images using the Fast Level Set Transform. Two conference papers were accepted (Gretsi). A journal paper has been accepted to IEEE IP. 3) Statistical analysis of visual processes (Coll. M. Lindenbaum, Technion) In the framework of the corresponding e-team, Y. Gousseau visited M. Lindenbaum for a week. Both theoretic and experimental work has been conducted at Technion, in particular concerning part-based object recognition.

3.17 Contribution by MTA SZTAKI

Researchers involved

S. Fazekas and D. Chetverikov (MTA SZTAKI), T.Amiaz and N.Kiryati (TAU-Visual), B.U.Toreyin, Y.Dedeoglu (Bilkent)

Activities & Achievements

A joint dynamic showcase demo created by SZTAKI, TAU-Visual and Bilkent. The demo and its description submitted and accepted for presentation at the MUSCLE demo session of CIVR 2007, Amsterdam. Work on developing a novel approach to background modelling based on dynamic texture started by SZTAKI. First promising results obtained. The approach is capable of modelling periodically moving and other dynamic backgrounds, such as escalators.

Problems

Lack of video data with strong dynamic backgrounds. Solution in progress: Creating such data and, possibly, adding it to the MUSCLE database DynTex.

Events

A talk on the current state of the joint MUSCLE dynamic texture research at MTA SZTAKI and TAU-Visual delivered by Chetverikov at the MUSCLE session of Fifth International Workshop on Content-Based Multimedia Indexing (CBMI), Bordeaux, 2007.

Publications

1. S. Fazekas and D. Chetverikov, "Dynamic texture recognition using optical flow features and temporal periodicity", Proc. Fifth International Workshop on Content-Based Multimedia Indexing (CBMI), Bordeaux, pp. 25-32, 2007.
2. S. Fazekas and D. Chetverikov, "Analysis and performance evaluation of optical flow features for dynamic texture recognition", Special Issue of Signal Processing: Image Communication journal on Content-Based Multimedia Indexing, vol. 22/7-8, pp. 680-691, 2007.

4.5 Contribution by UNIS

Researchers involved

B. Goswami

Activities & Achievements

As a contribution to E-Team on Person Detection, Recognition and Tracking: We are building an expression recognition system using lip shapes defined by B-splines. We are testing the

performance of the system in comparison to a system based on a level-set approach that used the most significant contour deformations in the face as expression descriptors. A paper on this work is in the course of preparation.

Events

Attended 4th Biometrics Summer School, Alghero, Italy

4 Overview activities in WP4

4.1 Contribution by CEA LIST

Researchers involved

Gregory Grefenstette

Activities & Achievements

CEA LIST continued research on natural language models for image recognition. Planning for future cooperation (visits, workshops) E-team "Integration of structural and semantic models for multimedia metadata management", submission of plan to MUSCLE board. Further Eteam activity: production of joint paper "Integrating Multimedia Metadata Management Systems", submitted to SAMT'2007 Organisation and participation in the MUSCLE sponsored international conference "Large Scale Semantic Access to Content (Text, Image, Video and Sound)" RIAO 2007, held at Carnegie Mellon University, May 30- June 1, 2007

Events

MUSCLE sponsored international conference "Large Scale Semantic Access to Content (Text, Image, Video and Sound)" RIAO 2007, held at Carnegie Mellon University, May 30- June 1, 2007

4.2 Contribution by TAU speech

Researchers involved

PIs: Arie Yeredor, David Burshtein

Activities & Achievements

1) We have continued our work on support vector machine training for improved hidden Markov modelling by designing an algorithm that uses soft segmentation data (via an N-best

list) and tested its performance. We have also started working on extending our new rescoring algorithm to connected speech. This work is also related to WP6.

2) We continued to pursue the single-channel audio separation problem, by exploring possible modifications to the algorithm proposed by Pearlmutter and Olsson in "Linear Program Differentiation for Single Channel Separation" (proceedings of MLSP'06). We explored possible modification to the authors' proposal of using Linear Programming (LP) solvers in order to represent a vector as sparsely as possible in an over-complete basis, using a gradient-based method. Several LP approaches were examined and implemented in Matlab, to enable comparison of the algorithms.

Publications

1. A. Sloin and D. Burshtein, "Support Vector Machine Training for Improved Hidden Markov Modeling," accepted for publication in IEEE Trans. on Signal Processing (accepted in May 2007).

2. H. Aronowitz and D. Burshtein, "Efficient Speaker Recognition Using Approximated Cross Entropy (ACE)," accepted for publication in IEEE Transactions on Audio, Speech and Language Processing, to be published September 2007 (accepted in May 2007).

4.3 Contribution by TU Vienna-IFS

Researchers involved

Jakob Frank, Thomas Lidy, Andreas Rauber

Activities & Achievements

The web service for Audio Feature Extraction was improved and thoroughly tested. This web service is now being extended and complemented by a web service providing the possibility for users to train and create Self-Organising Music Maps online. Effort is being put on integrating both services into an all-in-one web service for the creation of interactive Music Maps for access to music archives. Our paper on "Improving Genre Classification By Combination Of Audio And Symbolic Descriptors Using A Transcription System", written jointly with a research group from Alicante, Spain, was accepted for presentation at ISMIR 2007. A paper on "Bringing mobile based map access to digital audio to the end user" was accepted for presentation at the 14th International Conference on Image Analysis and Processing and the International Workshop On Visual And Multimedia Digital Libraries (VMDL07). Further, a paper on the joint analysis of audio and textual content of music (Multi-modal music information retrieval - visualisation and evaluation of clusterings by both audio and lyrics) was accepted and presented at the MUSCLE sponsored international conference "Large Scale Semantic Access to Content (Text, Image, Video and Sound)" RIAO 2007, held at Carnegie Mellon University, May 30- June 1, 2007. We started preparations for our submission to the Music Information Retrieval Evaluation eXchange (MIREX) 2007, where we will participate with the combined audio and symbolic music classification approach in several tasks.

Events

8th Conference Recherche d'Information Assistée par Ordinateur (RIAO'07), Pittsburgh, PA, USA, May 29th - June 1 2007 International Workshop On Visual And Multimedia Digital Libraries (VMDL07), Modena, Italy, September 13-14, 2007 MIREX 2007, Vienna, Austria, September 23 -27, 2007 International Conference on Music Information Retrieval, Vienna, Austria, September 23 -27, 2007

Publications

Thomas Lidy, Andreas Rauber, Antonio Pertusa, Jose Manuel Inesta. "Improving Genre Classification By Combination Of Audio And Symbolic Descriptors Using A Transcription System". International Conference on Music Information Retrieval 2007. Submitted. Robert Neumayer, Jakob Frank, Peter Hlavac, Thomas Lidy, and Andreas Rauber. Bringing mobile based map access to digital audio to the end user. In Proceedings of the 14th International Conference on Image Analysis and Processing (ICIAP'07), Modena, Italy, September 10 - 13 2007. Accepted for publication.

4.4 Contribution by AUTH

Researchers involved

C. Kotropoulos, M. Sedaaghi, E. Benetos, M. Kotti, G. Almpanidis

Activities & Achievements

The emotional state of an utterance is recognized using statistics of pitch, energy, and formant contours as features. The utterances and the features are arranged in the rows and the columns of the pattern matrix (also known as data matrix), respectively. By treating the columns of the pattern matrix as one-dimensional signals, we propose to smooth the measurements of each feature with either morphological filters or wavelets, before emotional speech classification is performed by the Bayes classifier that operates on a feature subset determined by the sequential floating forward selection algorithm on the grounds of the correct classification rate. The probability density function of the selected features within the Bayes classifier is assumed to be a multivariate Gaussian. The experimental results on the Danish Emotional Speech database indicate that the correct emotional speech classification rate could be as high as 84.5 % by applying proper smoothing. The achieved rate is the second best reported rate among 8 classifiers applied to the same database. This rate is obtained when the label information is considered. The correct classification rate reduces to 60.3% when the label information is ignored during pattern matrix smoothing. In this case, the reported rate is approximately 12% higher than our previously published rate on the same database, a fact that further supports the claim that smoothing of the pattern matrix is necessary before classification. Work has been done on gender classification by comparing different classifier accuracies on two emotional speech datasets. The following classifiers are compared: the universal gender models (UGMs), the probabilistic neural networks (PNNs), the support vector machines (SVMs), and the K-nearest neighbours (KNNs). The databases used are the Danish Emotional Speech (DES) database and the Berlin database of Emotional Speech. Numerous features are extracted for gender classification. Starting with an initial set of just 90 features and a maximum gender classification accuracy equal to 95% on DES database, a final

set of 1379 features is created and the best gender classification accuracy is 100%, when selecting 90 out of the 1379 features for the Berlin database of Emotional Speech. Moreover, different training/test sets are applied to investigate classifiers' tolerance to the amount of training data. Speech segmentation at a phone level imposes high resolution requirements in the short-time analysis of the audio signal. In this work, we employ the Bayesian information criterion corrected for small samples and model speech samples with the generalised Gamma distribution, which offers a more efficient parametric characterisation of speech in the frequency domain than the Gaussian distribution. Using a computationally inexpensive maximum likelihood approach for parameter estimation, we attest that the proposed adjustments yield significant performance improvement in noisy environments. Automatic musical instrument identification using a variety of classifiers is addressed. Experiments are performed on a large set of recordings that stem from 20 instrument classes. Several features from general audio data classification applications as well as MPEG-7 descriptors are measured for 1000 recordings. Branch-and-bound feature selection is applied in order to select the most discriminating features for instrument classification. The first classifier is based on non-negative matrix factorization (NMF) techniques, where training is performed for each audio class individually. A novel NMF testing method is proposed, where each recording is projected onto several training matrices, which have been Gram-Schmidt orthogonalized. Several NMF variants are utilized besides the standard NMF method, such as the local NMF and the sparse NMF. In addition, 3-layered multilayer perceptions, normalized Gaussian radial basis function networks, and support vector machines employing a polynomial kernel have also been tested as classifiers. The classification accuracy is high, ranging from 88.7% to 95.3%, thus outperforming the state-of-the-art techniques tested in the aforementioned experiment.

Publications

G. Almpantidis and C. Kotropoulos, "Automatic phonemic segmentation using the Bayesian information criterion with generalized Gamma priors," in Proc. XV European Signal Processing Conf., accepted. E. Benetos, M. Kotti, and C. Kotropoulos, "Large-scale musical instrument identification," in Proc. 4th Sound and Music Computing Conference, accepted.

4.6 Contribution by CNR-ISTI

Researchers involved

Graziano Bertini, Vincenzo Di Salvo, Massimo Magrini, Andreas Rauber.

Activities & Achievements

- Continuation in testing of ARIA Dynamic Sound (Release 2), an algorithm for compressed music transient enhancement, using ISMIR mp3 tracks collection. A protocol for subjective evaluation with expert and non expert audience is in the works, in order to estimate the differences between processed audio tracks vs original tracks using different ARIA settings. - Test of an algorithm for real-time control of video effects in a multimedia environment (Pandora System), using pitch and RMS parameters extracted from clarinet sounds. Preparation of a paper related to this topic. - Continued development of a web-based environment for exchanging multi-track audio/musical signals between European secondary schools (MODEM EU project).

Problems

A quantitative measurement of music quality perception enhancement is difficult to estimate, since listening experience is related to sound attributes such as clarity, vivacity, dynamic range width, etc., whose evaluation is intrinsically subjective.

Publications

Paper “Real-time sound parameters tracking in a multimedia system” submitted to DSP Application Day 2007 conference, Milan Sept. 2007.

4.7 Contribution by IRIT-UPS

Researchers involved

Régine André-Obrecht, Hélène Lachambre, Khalid Daoudi, Andrey Temko, Julien Pinquier

Activities & Achievements

The collaboration with UPC (Barcelona, Spain) started in April, has been continued with the presence of Andrey Temko, a PhD student at UPC, joined IRIT for a period of 3 months (until July 2007). Sequence kernels developed at IRIT have been applied to the problem of specific acoustic events detection. About the study of the Speech/Music/Song discrimination, an effort is produced to understand some errors: for example why song is detected during some pure instrumental music segment and why song is not always detected.

5 Overview activities in WP5

5.1 Contribution by Bilkent University

Researchers involved

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

Activities & Achievements

A novel method to detect forest fire smoke in video is under development. Manned lookout posts are commonly installed in the forests all around the Europe and the world. Most of these posts have electricity. Surveillance cameras can be placed on to these surveillance towers to monitor the scene for possible forest fires and they can be used to monitor the progress of the fire from remote centers. Forest fire smoke monitored from these far away posts have different temporal and spatial characteristics than smoke monitored from a nearby surveillance camera. Smoke at far moves much slower than a nearby smoke. In addition, the

flicker within and on the boundaries of smoke regions disappear as smoke gets further away from the camera. Two separate background images are estimated by an IIR filter with different forgetting factors. Slow moving regions within the viewing range of the camera are revealed in the difference image between these background images. Smoke regions are smooth with little texture information and have little energy contained in high-frequency subbands. Possible smoke regions are analyzed spatially in wavelet domain.

5.2 Contribution by TU Vienna-IFS

Researchers involved

Robert Neumayer, Andreas Rauber

Activities & Achievements

We presented the paper "Multi-modal music information retrieval - visualisation and evaluation of clusterings by both audio and lyrics" at the 8th Conference Recherche d'Information Assistée par Ordinateur (RIAO'07), Pittsburgh, PA, USA, May 29 - June 1 2007. Research on clustering of music archives using multi-modal approaches combining audio features and textual information gathered from the web continued.

Events

8th Conference on Recherche d'Information Assistée par Ordinateur (RIAO'07), Pittsburgh, PA, USA, May 29 - June 1 2007.

Publications

Robert Neumayer and Andreas Rauber. Multi-modal music information retrieval - visualisation and evaluation of clusterings by both audio and lyrics. In Proceedings of the 8th Conference Recherche d'Information Assistée par Ordinateur (RIAO'07), Pittsburgh, PA, USA, May 29 - June 1 2007.

5.3 Contribution by AUTH

Researchers involved

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

Activities & Achievements

On-going research on speaker diarization within the showcase "Movie Summarization and Skimming Demonstration" has been continued aiming at dialogue detection in movie scenes. A speaker diarization system, that makes no assumptions on the number of speakers participating in the conversation, has been developed and tested, with respect to the overall classification error. In particular three hierarchical algorithms are utilized to produce the cluster ensemble, namely the average group linkage, the weighted average group linkage and

the Ward's hierarchical clustering method. The input of the hierarchical clustering algorithms is a distance matrix, created by calculating the distances among the speech segments, using the distance metric between two covariance matrices. In total, 30 different partitions are produced. For each partition, the number of clusters to be created and the clustering algorithm to be applied are randomly chosen. Next, the co-association matrix is computed. Each speech segment is considered to belong to the same cluster with another segment, when the respective entry of the co-association matrix exceeds a threshold. Two variants of the algorithm are available. The first variant does not cluster the speech segments that are considered to be outliers, while the second one does. The best overall classification error value measured is 14.78%, leading to the creation of 7 clusters, when outliers are not included in clustering. Generally, the algorithm tends to over-estimate the number of clusters. During the reporting period, the work towards a face clustering algorithm that utilizes a dissimilarity matrix built using SIFT image features continued. More specifically, the method was fine-tuned and experimentally tested in facial images detected in movies. Experimental evaluation involved three well known clustering validity measures in order to assess the performance of the clustering results, in an objective fashion. These measures were the F measure, the overall entropy and the Γ statistic. Results showed that the proposed method can perform very efficiently.

Publications

M. Kotti, E. Benetos, and C. Kotropoulos, "Neural network-based movie dialogue detection," in Proc. 2007IEEE Engineering Applications of Neural Networks, accepted.

5.4 Contribution by IRIT-UPS

Researchers involved

Julien Pinquier, Frédéric Gianni

Activities & Achievements

During this period, we have finalized the showcase ACADI (Automatic Character in Audiovisual Document Indexing): - management of the multi-track, - synchronization of the audio and video reading, - improvement of the visual interface. We have also proposed a showcase extension for JPA4: on one hand to improve fusion of the segmentations (speakers and costumes) and on the other hand to evolve the interface in an annotator, in order to produce ground truth. Our activity plan was accepted. About achievements, ACADI interface evolves, note in particular: - association of the audio and video reading to each segment, - simultaneous visualization of several tracks (speaker and costume detection for example), - addition of a reading cursor, - addition of colors to distinguish the various tracks and partner logos.

5.5 Contribution by ICCS-NTUA

Researchers involved

G. Papandreou, A. Katsamanis, V. Pitsikalis, P. Maragos (ICCS-NTUA)

Activities & Achievements

Audio-Visual Interaction for Speech Processing

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

Publications

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

5.6 Contribution by ICCS-NTUA

Researchers involved

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

Activities & Achievements

Real-time Audio-visual Automatic Speech Recognition Demonstrator Showcase

We have continued during the reporting period the effort to build a real-time audio-visual automatic speech recognition demonstrator, as part of the Muscle Showcasing initiative. We have spent significant effort in streamlining the visual front-end of the prototype. Trying to overcome the limitations of our previous research-level visual front-end module, which was designed for off-line feature extraction in well-defined scenarios, such as processing videos shot under carefully controlled conditions, we have made the following improvements: 1. We have integrated a fast adaboost-based face detector in our visual front-end. The role of the face detector is to initialize the face tracker at the first frame, and also whenever tracking fails. 2. We have made algorithmic advances in the techniques for AAM fitting, yielding efficient and accurate computer vision algorithms for parametric model fitting and tracking. 3. Repeated image resampling at updated shape warps (texture mapping in computer graphics terminology) constitutes a significant part of the computational load during model fitting. We build on advanced features of modern graphics cards (GPUs) and perform texture-mapping very efficiently on the GPU. We are currently working on integrating all these components into our prototype's visual front-end.

5.7 Contribution by ICCS-NTUA

Researchers involved

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

Activities & Achievements

Audiovisual Attention Modelling and Salient Event Detection

Although human perception appears to be automatic and unconscious, complex sensory mechanisms that form the pre-attentive component of human understanding and lead to awareness exist. Considerable research has been carried out into these pre-attentive 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 modelling, we extract attention curves using features around the spatio-temporal 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.

Publications

K. Rapantzikos, G. Evangelopoulos, P. Maragos, and Y. Avrithis, “An audiovisual saliency model for movie summarization”, accepted for publication, IEEE Workshop on Multimedia Signal Processing (MMSP-2007), 2007.

5.8 Contribution by ICCS-NTUA

Researchers involved

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

Activities & Achievements

Movie Summarization and Skimming Demonstrator Showcase

During the reporting period, we have launched an effort to build a Movie Summarization and Skimming Demonstrator, as part of the Muscle Showcasing initiative Participating partners are ICCS-NTUA (leader), TSI-TUC, AUTH, and INRIA-Textmex. 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.

5.9 Contribution by ICCS-NTUA

Researchers involved

Petros Maragos (ICCS-NTUA), Alexandros Potamianos (TSI-TUC) and Patrick Gros (INRIA-Textmex)

Activities & Achievements

Book on "Multimodal Processing and Interaction: Audio, Video, Text"

Petros Maragos (ICCS-NTUA), Alexandros Potamianos (TSI-TUC) and Patrick Gros (INRIA-Textmex), Editors

The book under preparation is covering the thematic areas of WP5 (former JPA2 WP6 & WP10). 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 WP5 members. The book proposal was submitted to Springer-Verlag and has been approved.

5.10 Contribution by ICCS-NTUA

Researchers involved

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

Activities & Achievements

Audio-Visual Speech Inversion

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

Publications

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

5.11 Contribution by INRIA - Texmex

Researchers involved

Xavier Naturel Patrick Gros

Activities & Achievements

Movie summarizer showcase

Concerning the movie summarizer showcase, INRIA's activity was to generate the movie summary based on input features from NTUA and AUTH. Our task has been to collect the data, audio-visual features from NTUA and speaker clustering results from AUTH, computed on the movie database previously defined by the partners. Using these features, the video summary is made using a two-pass K-means clustering. Two methods have been implemented: one fully automatic and one where the user is allowed to indicate his preferences by choosing a few clusters from the first clustering phase.

5.12 Contribution by VTT

Researchers involved

Sanni siltanen, mika Hakkarainen

Activities & Achievements

Development of the Showcase "Augmented Assembly Using a Multimodal Interface" demonstration. Improvement of the multimodal user interface concentrating on the gesture input and configuration of devices.

6 Overview activities in WP6

6.1 Contribution by Bilkent University

Researchers involved

B. Ugur Toreyin, R. Gokberk Cinbis, Yigithan Dedeoglu, A. Enis Cetin

Activities & Achievements

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 behaviour. IR image frames are also analyzed spatially.

Boundary 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 modelling and wavelet domain boundary modelling.

Publications

B. Uğur Töreşin, Ramazan Gökberk Cinbiş, Yiğithan Dedeođlu, A. Enis Çetin, "Fire Detection in Video using Wavelet Analysis", Optical Engineering, Vol. 46, No. 6, pp. 067204-1-067204-9, June 2006.

6.2 Contribution by TU Vienna-IFS

Researchers involved

Rudolf Mayer, Robert Neumayer, Thomas Lidy, Andreas Rauber

Activities & Achievements

Our research on visualizations which aid interpretation of clusterings on Self-Organizing Maps resulted in three novel visualization methods: A Thematic Class Distribution Visualization, a Sky-Metaphor visualization and the metro visualization. All three novel visualization methods are to be publicised in research papers, which were already accepted for publication, the metro visualisation at the 6th International Workshop on Self-Organizing Maps (WSOM'07), the Sky-Metaphor Visualisation at the 7th International Conference on Knowledge Management, and the Thematic Class Distribution at the International Conference on Artificial Neural Networks (ICANN'07). Moreover, we investigated a novel classification method, which we call "Decision Manifolds" and which is inspired by Self-Organization. Experimental results with this novel kind of classifier are described in a paper, which was accepted for publication at the 6th International Workshop on Self-Organizing Maps (WSOM'07). Work continued on the e-Team on Active and Semi-Supervised Learning and an e-Team meeting was held in Vienna with Michel Crucianu/INRIA-IMEDIA and Rudolf Mayer/TU Vienna-IFS between May 21 - 23, 2007.

Events

e-Team Meeting on Active and Semi-Supervised Learning, Vienna, Austria, May 21 - 23, 2007
6th International Workshop on Self-Organizing Maps (WSOM'07), Bielefeld, Germany, September 3 - 6 2007.

7th International Conference on Knowledge Management (I-KNOW'07), Graz, Austria, September 5 - 7 2007.
International Conference on Artificial Neural Networks (ICANN'07), Porto, Portugal, September 9 - 13 2007.

Publications

Robert Neumayer, Rudolf Mayer, and Andreas Rauber. Component selection for the metro visualisation of the SOM. In Proceedings of the 6th International Workshop on Self-Organizing Maps (WSOM'07), Bielefeld, Germany, September 3 - 6 2007.

Georg Pözlbauer, Thomas Lidy and Andreas Rauber: Decision Manifolds: Classification Inspired by Self-Organization. Accepted for publication at the 6th International Workshop on Self-Organizing Maps, Bielefeld, Germany, September 3-6, 2007.

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

Rudolf Mayer, Taha Abdel Aziz, and Andreas Rauber. Visualising Class Distribution on Self-Organising Maps. In Proceedings of the International Conference on Artificial Neural Networks (ICANN'07), Porto, Portugal, September 9 - 13 2007. Springer Verlag.

6.3 Contribution by TUG

Researchers involved

Amir R. Saffari A. A., Michael Grabner, Helmut Grabner, Joachim Pehserl, Petra Korica-Pehserl, Horst Bischof

Activities & Achievements

1. We are currently experimenting the effect of different clustering algorithms together with those developed by ourselves for visual object recognition in a bag-of-words framework for large-scale datasets.
2. We developed an approach which integrates detection, recognition and tracking by formulating all tasks as binary classification problems. Because of its efficiency it is well suited for robots or other systems with limited resources but nevertheless demonstrates robustness and comparable results to state-of-the-art approaches. We use a common over-complete representation which is shared by the different modules. By means of the integral data structure an efficient feature computation is performed enabling the usage of this system for real-time applications.

Publications

Michael Grabner, Helmut Grabner, Joachim Pehserl, Petra Korica-Pehserl, Horst Bischof "Flea, do you remember me?" Asian Conference on Computer Vision 2007 (ACCV'07), to appear

6.4 Contribution by AUTH

Researchers involved

C. Kotropoulos, D. Ververidis, M. Sedaaghi

Activities & Achievements

Our ongoing research on feature selection using has led to 2 conference papers accepted during this reporting period. An accurate estimate of the cross-validated prediction error variance in Bayes classifiers has been derived. Self-adaptive genetic algorithms have been employed to search for the worst performing features with respect to the probability of correct classification achieved by the Bayes classifier in a first stage. That is, a genetic algorithm-based implementation of backward feature selection has been developed. These features are subsequently excluded from sequential floating feature selection employing the probability of correct classification achieved by the Bayes classifier as criterion. In a second stage, self-adaptive genetic algorithms are employed to search for the worst performing utterances with respect to the same criterion.

Publications

M. Sedaaghi, C. Kotropoulos, and D. Ververidis, "Improving speech emotion recognition using adaptive genetic algorithms," in Proc. XV European Signal Processing Conf., accepted. D. Ververidis and C. Kotropoulos, "Accurate estimate of the cross-validated prediction error variance in Bayes classifiers", in Proc. 2007 IEEE Machine Learning in Signal Processing, accepted.

6.5 Contribution by ISTI-CNR

Researchers involved

Luigi Bedini, Sara Colantonio, Ercan Kuruoglu, Davide Moroni, Emanuele Salerno, Ovidio Salvetti, Anna Tonazzini

Activities & Achievements

The activities of E-team "Unsupervised image segmentation" are continuing by testing the statistical data model developed for blind source separation applied to astrophysical image analysis. A first version of the derived MCMC strategy is now coded in Matlab (R) and is being tested at UCD with artificially generated data. The data model will be further enhanced after assessing the results obtained by the present tests and other experiments with more realistic data, taken from the sky maps simulated by the astrophysicists in the ESA's "Planck" cosmological mission. Also, research work continues on the alternative approaches to blind separation studied at ISTI, namely, use of second-order statistics and spatial correlation ("Correlated component analysis", with application to astrophysical imaging), and local maximization of nongaussianity ("MaxNG", with application to remote-sensed hyperspectral data). The results from both approaches will be presented at the forthcoming CLeMUS (Computational Learning Methods for Unsupervised Segmentation) session of the KES 2007 Conference.

Within the frame of integrating inferential and computational reasoning for supporting clinical decision making, an ontological model of the cytology domain has been developed as a basis for the feature-based analysis of microscopic cell images. More precisely, a Cell Image Analysis Ontology – CIAO – has been defined for the cell image features domain. CIAO is a structured ontology that relates one another, cells or cell parts, microscopic images and

cytometric features. Such ontology has been integrated in a software package, equipped with cell segmentation methods and a library of algorithms for features computation, which supports biologists and clinicians in their analysis processes. The final aim is the development of an automated diagnosis system. The results of the activity have been detailed in a paper accepted to the Open German Russian Workshop 2007.

Events

Emanuele Salerno attended the "Planck" Consortium plenary meeting in Toulouse, 18-20 June 2007.

Publications

S. Colantonio, I.B. Gurevich, M. Martinelli, O. Salvetti, Yu. Trusova. Ontology driven Approach to Cell Image Analysis. Accepted to Open German Russian Workshop, Ettlingen, Germany, 20-23 August 2007.

6.6 Contribution by UCD

Researchers involved

Pádraig Cunningham Derek Greene Ken Bryan Anton Zamolotskikh

Activities & Achievements

Pádraig Cunningham attended the Muscle review in Nice from 2nd to 4th May. Ken Bryan and Pádraig Cunningham continued work on the Muscle book on ML and Multimedia to be published by Springer. The main research activities in the UCD ML group that relate to Muscle are work on Clustering and Dimension Reduction. In clustering we are working on semi-supervised clustering and the problem of identifying useful constraints. The approach we are evaluating at the moment entails an ensemble approach to this problem. We are also working on the application of dimension reduction techniques to the specialised problem of dimension reduction of metabolomics data.

Events

Anton Zamolotskikh attended the 5th IEEE International Workshop on Content-Based Multimedia Indexing in France in June.

Publications

The following paper was accepted for publication at ECML 2007: Greene, D., Cunningham, P., (2007) Constraint Selection by Committee: An Ensemble Approach to Identifying Informative Constraints for Semi-Supervised Clustering, 18th European Conference on Machine Learning (ECML 07), J. N. Kok, J. Koronacki, R. Lopez de Mantaras, S. Matwin, D. Mladenič, A. Skowron (Eds.), Lecture Notes in Artificial Intelligence 4701, Springer 2007. The following paper was published at the IEEE workshop on Content-Based Multimedia Indexing which is sponsored by the Muscle NoE. Zamolotskikh, A., Cunningham P. (2007) An assessment of alternative strategies for constructing EMD-based kernel functions for use

in an SVM for image classification, 5th IEEE International Workshop on Content-Based Multimedia Indexing, pp11-17, IEEE Press, E. Pauwels (ed.) 2007.

6.7 Contribution by UPMC

Researchers involved

Matthieu Cord, David Picard, Julien Gony

Activities & Achievements

We made research activities on active learning for content-based image and text database retrieval, and carried out a new demonstrator RETIN to process text + image queries and search in multimedia database in collaboration with ENSEA partner. We continued the Collaboration with F. Precioso (ENSEA partner) on video content analysis focused on face detection. Matthieu Cord continued the editing work with P. Cunningham (UCD partner) on book on ML for Multimedia to be published by Springer. We prepared a paper on "Distributed Content Based Image Retrieval using Ant Algorithm" submitted to IEEE trans on Multimedia application journal

Events

Matthieu Cord participated to the Fifth International Workshop on Content-Based Multimedia Indexing (CBMI 2007) with special MUSCLE event.

Publications

D. Gorisse, M. Cord, M. Jordan, S. Philipp-Foliguet, F. Precioso, 3D Content-based retrieval in artwork databases, 3DTV Conference, 2007 C. Iovan, D. Boldo, M. Cord and M. Erikson, Automatic Extraction and Classification of Vegetation Areas from High Resolution Images in Urban Areas, SCIA conference, 2007 C. Iovan, D. Boldo, M. Cord, Automatic Extraction of urban vegetation structures from High Resolution Imagery and digital elevation models, Urban remote sensing joint event, Paris, 2007

6.8 Contribution by IRIT-UPS

Researchers involved

Khalid Daoudi

Activities & Achievements

During these two months, the four laboratories (CMM, CWI, INRIA Texmex and IRIT) have prepared an extension of the e-team "Dynamic Kernels" under the responsibility of Khalid Daoudi. The proposition may be summarized into two main points: - An important part of the activity will be dedicated to extend the work done on classification to regression, from both theoretical and practical point of view with an application to the indexing of audio/video

sequences. - The second important part will be dedicated to the continuation of collaboration between IRIT and UPC on the problem of acoustic events detection using SVMs (see WP4).

Publications

Feature space Mahalanobis sequence kernels: Application to SVM Speaker Verification. J. Louradour, K. Daoudi and F. Bach IEEE Transactions on Speech, Audio and Language Processing. To appear (accepted in June 2007).

6.9 Contribution by INRIA Ariana

Researchers involved

Josiane Zerubia, Ian Jermyn, Ting Peng.

Activities & Achievements

Ting Peng, joint PhD student of INRIA Ariana and the LIAMA Institute in Beijing has created phase field versions of the nonlinear models described under WP3, and derived their functional derivatives. These are considerably more complex than the linear case. Nevertheless, Ms Peng has found a way to implement them just as efficiently, thus rendering feasible the gradient descent algorithm for the new model.

Publications

A paper has been accepted at the conference BMVC 2007 (Peng).

6.10 Contribution by ENSEA

Researchers involved

Dr. P.-H. Gosselin, J. Lebrun, Prof. S. Philipp-Foliguet (ENSEA)

Activities & Achievements

We are currently intending to extend our object class retrieval system using kernels on bags to kernels on graphs. Indeed, we want to consider graph representation in order to, for example, include spatial constraints between the fuzzy regions obtained after segmentation process of images to be classified. First results with pairs of fuzzy regions are very encouraging. We made the state of art in this domain and have carried out experiments with several techniques using graphs of fuzzy regions or points of interest. We also extended the RETIN retrieval system to text + image retrieval, which shows its efficiency in the Imavegal live contest.

7 Overview activities in WP7

7.1 Contribution by Bilkent University

Researchers involved

Behcet Ugur Toreyin, Yigithan Dedeoglu, and A. Enis Cetin

Activities & Achievements

We study detection and segmentation of dynamic textures in participation with our E-team: "Dynamic Textures and Adaptive Background Modelling". Dynamic textures are common in natural scenes. Examples of dynamic textures in video include fire, smoke, clouds, trees in the wind, sky, sea and ocean waves etc. In this showcase, (i) we develop real-time dynamic texture detection methods in video and (ii) present solutions to video object classification based on motion information.

Events

MUSCLE participated to technical demos at ACM CIVR 2007. Dynamic Texture, Detection, Segmentation and Analysis is presented as a technical demonstration at this event. The 6th International Conference on Image and Video Retrieval (CIVR 2007) took place in Amsterdam, the Netherlands during 9-11 July 2007. The showcase events are sponsored by MUSCLE NoE.

Publications

Behcet Toreyin, Yigithan Dedeoglu, Enis Çetin, Sándor Fazekas, Dmitry Chetverikov, Tomer Amiaz, Nahum Kiryati, "Dynamic Texture Detection, Segmentation and Analysis", Proceedings of ACM Conference on Image and Video Retrieval, June 2007.

7.2 Contribution by ACV

Researchers involved

Csaba Beleznai

Activities & Achievements

As a participant of the showcase "Unusual Event Detection" we worked on documentation to support the integration of ACV algorithmic modules into the showcase framework. In addition we provided detailed description of the algorithmic modules which has been incorporated into a publication submitted to the CIVR 2007 (July 9-11, 2007). We performed evaluation of our algorithmic components on some showcase video data.

Events

The demo Flexible Test-bed for Unusual Behavior Detection by István Petrás (Computer and Automation Research Institute, Hungarian Academy of Sciences, Hungary); Levente Kovács (Computer and Automation Research Institute, Hungarian Academy of Sciences, Hungary); Behcet Toreyin (Bilkent University, Turkey); Csaba Beleznai (ACV, Austria); Zoltan Szlavik (Mta-SZTAKI, Hungary); Ugur Gudukbay (Bilkent University, Turkey); Yigithan Dedeoglu (Bilkent University, Turkey); László Havasi (Peter Pazmany Catholic University, Hungary); Enis Cetin (Bilkent University, Ankara, Turkey); Montse Pardas (Technical University of Catalonia, Spain); Tamas Sziranyi (Computer and Automation Research Institute of the Hungarian Academy of Sciences, Hungary); Cristian Canton (Universitat Politecnica de Catalunya, Spain) will be demonstrated at CIVR 2007. The paper of the demo will also be presented at CIVR 2007 in July.

7.3 Contribution by Bilkent University

Researchers involved

Yigithan Dedeoglu, Behcet Ugur Toreyin, Ugur Gudukbay, and A. Enis Cetin

Activities & Achievements

Visual surveillance and activity analysis is an active research field of computer vision. As a result, there are several different algorithms produced for this purpose. To obtain more robust systems, it is desirable to integrate the different algorithms. To help achieve this goal, we propose a flexible, distributed software collaboration framework and present a prototype system for automatic event analysis. We provide a transparent and distributed architecture for easy integration of third party modules into a common framework to facilitate easier research collaboration and evaluation. The setup is hierarchical thus helping the scalability of the whole framework. The actual implementation integrates diverse algorithms forming a test-bed for unusual activity detection. Various complex surveillance related algorithms, such as human and body action, tracking and motion activity algorithms are integrated into one system. In the case of detecting unusual motion occurrences, we refer to the term unusual in statistical sense.

Events

MUSCLE participated to technical demos at ACM CIVR 2007. Flexible Test-bed for Unusual Behavior Detection is presented as a technical demonstration at this event. The 6th International Conference on Image and Video Retrieval (CIVR 2007) took place in Amsterdam, the Netherlands during 9-11 July 2007. The showcase events are sponsored by MUSCLE NoE.

Publications

István Petrás, Levente Kovács, Behcet Toreyin, Csaba Beleznai, Zoltan Szlavik, Ugur Gudukbay, Yigithan Dedeoglu, László Havasi, Enis Cetin, Montse Pardas, Tamas Sziranyi, Cristian Canton, "Flexible Test-bed for Unusual Behavior Detection", Proceedings of ACM Conference on Image and Video Retrieval (CIVR), June 2007.

7.4 Contribution by Bilkent University

Researchers involved

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

Activities & Achievements

We are currently preparing our showcases for participation to IBC 2007 event which will be held on 7-11 September 2007, in Amsterdam. The aim of participation is to demonstrate the current results of MUSCLE through showcase demos. Showcases that are carried out under WP7 will be demonstrated in this exhibition. It is planned to have 3-4 demos each day simultaneously. The stand area and the furniture design for the exhibition is arranged. For disseminating the results of the project, demo flyers will be distributed. Project and demo posters will be prepared for this exhibition. Each contributing partner will also bring additional dissemination material about their demo.

Events

IBC 2007 event, which will be held on 7-11 September 2007, in Amsterdam.

7.5 Contribution by UPC

Researchers involved

Montse Pardas, Dafnis Batalle

Activities & Achievements

An events detection module based on data produced by the body and motion analysis has been developed. A vector of features is extracted for classification, including axis of the ellipse fitted to the detected foreground object, orientation of this ellipse, and Hu moments of the Motion Histori Image and Motion Energy Image. Dafnis Batallé, from UPC has visited Sztaki for one week within the context of this showcase. During this stay the tools developed at UPC for Human model and motion based unusual event detection have been integrated in the Showcase platform. These tools have also been adapted and trained to detect the following events: fights, pick up an object and sit down.

7.6 Contribution by TAU-VISUAL

Researchers involved

Tomer Amiaz and Nahum Kiryati (TAU-VISUAL). Collaboration with Sandor Fazekas and Dmitry Chetverikov (MTA-SZTAKI).

Activities & Achievements

During this period, we focused on the development of a graph-cut based dynamic texture segmentation algorithm. It is intended to accelerate our earlier algorithm, which is based on variational principles. Dynamic texture segmentation is an interesting way to detect important phenomena such as smoke in video. The computational acceleration that we work on is necessary for real-time operation, which is crucial for actual applications.

Events

The publication called "Detecting Regions of Dynamic Texture" is presented at First International Conference on Scale Space Methods and Variational Methods in Computer Vision (SSVM 2007) which was held at Ischia, Italy between May 30 - June 2, 2007.

Publications

T. Amiaz, S. Fazekas, D. Chetverikov and N. Kiryati, "Detecting Regions of Dynamic Texture", Proc. First International Conference on Scale Space Methods and Variational Methods in Computer Vision (SSVM 2007), Ischia, Italy, May 30 - June 2, 2007.

7.7 Contribution by TU Vienna-IFS

Researchers involved

Rudolf Mayer, Robert Neumayer, Thomas Lidy, Andreas Rauber

Activities & Achievements

Major work was done for the Book Chapter on "Multi-Modal Analysis of Text and Audio Features for Music Information Retrieval" for the MUSCLE Book on Multimodal Processing and Interaction. The Book Chapter "A Synthetic 3D Multimedia Environment", jointly written with the MUSCLE partners from EC3 within the showcase project "Shaping 3-dimensional Multimedia Environments", which was accepted for publication in "Computational Intelligence in Multimedia Processing: Recent Advances", Series "Studies in Computational Intelligence", Springer Verlag, was revised.

Events

Shaping 3-dimensional Multimedia Environments demo will be demonstrated at CIVR 2007.

Publications

Ronald Genswaidner, Helmut Berger, Michael Dittenbach, Andreas Pesenhofer, Dieter Merkl, Andreas Rauber and Thomas Lidy: "A Synthetic 3D Multimedia Environment". Accepted for publication in "Computational Intelligence in Multimedia Processing: Recent Advances", Series in "Studies in Computational Intelligence", Springer Verlag.

7.8 Contribution by TU VIENNA-PRIP

Researchers involved

Allan Hanbury, Branislav Micusik, Julian Stöttinger

Activities & Achievements

Work has been done on the two showcases in which we are involved.

For the evaluation showcase, the queries were created, consisting of nine text queries and nine visual queries. Examples of the two query types are given below:

- Text queries: "Find images of snowy mountains", "Find images of hotels with swimming pools", "Find images with three people on a beach".
- Visual queries: "What is the name of the monument in the provided image?" (i.e. find other images of the same monument and look at the text annotation).

The ground truth for each image was also created. All 21000 images in the dataset were examined manually to find those satisfying each of the queries. The final organisation of the live image retrieval event at the CIVR was completed.

For the object recognition showcase, the conversion of the code for colour interest point detection from MATLAB to C++ was completed. A web interface for the object recognition application was also created. The system was brought up to the stage allowing it to be demonstrated at the demo session of the CIVR. As part of the activity, Julian Stöttinger spent one month at the INRIA-IMEDIA.

Events

The Object Recognition and the image retrieval showcases will be demonstrated at CIVR 2007.

7.9 Contribution by Technion - ML

Researchers involved

Saher Esmeir, Michael Lindenbaum, Shaul Markovitch

Activities & Achievements

We have been working in this period mainly on cost-sensitive e-learning. Real-world usage of Machine learning techniques for multi-media applications, often involve nonuniform testing costs and misclassification costs. As the complexity of these applications grows, the management of resources during the learning and classification processes becomes a challenging task. We developed ACT (Anytime Cost-sensitive Trees), a novel framework for operating in such environments. ACT is an anytime algorithm that allows trading computation time for lower classification costs. It builds a tree top-down and exploits additional time

resources to obtain better estimations for the utility of the different candidate splits. Using sampling techniques ACT approximates for each candidate split the cost of the subtree under it and favours the one with a minimal cost. Given a tree, we estimate the expected costs when classifying a future case. This cost consists of two components: the test cost and misclassification cost. Assuming that the distribution of future cases would be similar to that of the learning examples, we can estimate the test costs using the training data. Given a tree, we calculate the average test cost of the training examples and use it to approximate the test cost of new cases. The misclassification cost is estimated using the expected error (in the same way C4.5 used it for pruning). Due to its stochastic nature ACT is expected to be able to escape local minima, into which greedy methods may be trapped. Our framework has 4 major advantages:

- (1) it uses a non-greedy approach to build a decision tree and therefore is able to overcome local minima problems,
- (2) it evaluates entire trees and therefore can be adjusted to any cost scheme that is defined over trees.
- (3) it exhibits good anytime behaviour and produces significantly better trees when more time is available, and
- (4) it can be easily parallelized and hence can benefit from distributed computer power.

To evaluate ACT we have designed an extensive set of experiments with a wide range of costs. The experimental results show that ACT is superior over ICET and EG2. Significance tests found the differences to be statistically strong. ACT also exhibited good anytime behaviour: with the increase in time allocation, there was a decrease in the cost of the learned models. ACT is a contract anytime algorithm that requires its sample size to be pre-determined.

In addition to the above activity, we worked on preparing the final version and presenting a paper on specular object reconstruction in CVPR2007.

We also worked on preparing the final version and presenting a paper on saliency algorithm design, in a special session dedicated to MUSCLE activity in the International workshop on Content Based Multimedia Indexing (CBMI), held in Bordeaux. This was an opportunity to meet many of the MUSCLE NoE and to hear about their work.

A meeting on the E-team: Statistical Analysis of Visual Processes, was held in Israel. For about a week, we met many times and discussed mostly two issues: analysis of part based recognition, and high level image modelling.

Publications

Saher Esmeir and Shaul Markovitch, Anytime Induction of Cost-sensitive Decision Trees, was submitted (and now accepted) to NIPS 2007. S. Rozenfeld, I. Shimshoni, and M. Lindenbaum. Dense mirroring surface recovery from 1D homographies and sparse correspondences. In CVPR, 2007. R. Golubchyc and M. Lindenbaum. The analysis of saliency processes and its application to grouping cues design. In Int. workshop on Content Based Multimedia Indexing, pages 18-24 2007.

7.10 Contribution by KTH

Researchers involved

Björn Granström, Olov Engwall, Preben Wik, David House

Activities & Achievements

Much work has been devoted to realising/implementing the Muscle showcase "Articulatory talking Head" that will be on display during the MUSCLE showcase demo at CIVR 2007 in Amsterdam. The Showcase will be extended and improved under a continuation grant during the second half of 2007. The extension of the showcase aims at displaying extra information on/with the talking head. First, some prosody information will be displayed. Prosody is a key in perception of speech and is one of the most difficult components to acquire in learning a foreign language or during speech therapy for deaf people. Second, the talking head will display information about the speech production not visible as articulations. Also experiments have been run on how auditory and visual cues for speaker friendliness can be modelled in multimodal speech synthesis. This work scientifically also falls within WP5.

Events

Articulatory Talking Head demo will be demonstrated at CIVR 2007.

Publications

Granström, B., & House, D. (2007). Inside out - Acoustic and visual aspects of verbal and non-verbal communication (Keynote Paper). Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken, 11-18. House, D. (2007). Integrating Audio and Visual Cues for Speaker Friendliness in Multimodal Speech Synthesis. Interspeech 2007.

7.11 Contribution by IRIT-UPS

Researchers involved

Julien Pinquier, Frédéric Gianni, Thomas Foures, Philippe Joly

Activities & Achievements

The showcase ACADI (Automatic Character in Audiovisual Document Indexing) has been finalized in June (see scientific complement in WP4 report). The evaluation portal of the CASEWP showcase has been finalized by the end of June. The evaluation tool has been integrated as a set of services on the NPDS portal generator. The three services are: results format checking, results evaluation submission and evaluation graphical synthesis. This last service includes an historic manager able to handle all the previously submitted results. Several possible extensions have been identified such as new metrics integration, the policy definition for anonymous result submission, and the improvement of the automatic historic manager.

Events

Participation to Large-Scale Semantic Access to Content (Text, Image, Video and Sound) – RIAO'2007 in Pittsburgh (USA) on May 30 and June 1 2007: presentation of the ACADI interface.

7.12 Contribution by MTA SZTAKI, Bilkent University, Politecnica de Catalunya, ACV

Researchers involved

István Petrás, Csaba Beleznai, Yiğithan Dedeoğlu, Montse Pardàs, Levente Kovács, Zoltán Szlávik, László Havasi, Tamás Szirányi, B. Uğur Töreyn, Uğur Gündükbay, A. Enis Çetin, Cristian Canton-Ferrer

Activities & Achievements

FLEXIBLE TEST-BED FOR UNUSUAL BEHAVIOR DETECTION

We made cooperations in research and showcase with involved institutes and, visits to common work Showcase demo system: The architecture according to the current trend and software tools is as flexible as possible. The modules can be distributed over the network; they are organized into a hierarchical structure. The structure can be separated into four main entities: the client's web interface, the server (possibly but not necessarily including the web server) the controller, and the communication inter-face embedded into the user module. Each component operates autonomously communication through RPC requests over TCP/IP.

a) The web interface gives the user transparent access to the control of the modules and can display their output. It is written in XHTML/Javascript and capable of displaying MJPEG streams.

b) The server is the central element of the system; it sits on the top of the hierarchy. It delegates the tasks to the modules and coordinates the execution of the different modules: starts and stops modules and sets their parameters. The server can have more than one controller.

c) The controller is a mid-level entity. Its role is to serve as a gateway between the modules on the local network. One controller can have several modules connected to it.

d) The communication interface seamlessly integrates into the user module. It is implemented as a C++ class. This makes it possible not to interfere with the existing user code. Through the interface the modules can transparently exchange images frames and other data.

One advantage of the architecture is that there is no need to publicly provide any proprietary source code, yet it is still possible to integrate heterogeneous modules through TCP/IP.

Problems

Visual surveillance and activity analysis is an active research field of computer vision. As a result, there are several different algorithms produced for this purpose. To gain more robust

systems it is desirable to integrate the different algorithms. To help this we propose a flexible, distributed software framework and present a prototype system for automatic event analysis.

Events

Unusual behaviour Detection demo will be demonstrated at CIVR 2007.

7.13 Contribution by INRIA Ariana

Researchers involved

Josiane Zerubia, Ian Jermyn, Avik Bhattacharya, Peter Horvath.

Activities & Achievements

Avik Bhattacharya attended CBMI 2007 to present his paper. Ian Jermyn and Peter Horvath presented the work done within MUSCLE to a number of visitors to the Ariana project during May and June. In particular, Ariana's work within MUSCLE was presented by Josiane Zerubia and Ian Jermyn to Tom Wassenar from the Joint Research Council of the European Union. This has resulted in collaborative work and a software transfer.

7.14 Contribution by IBAI

Researchers involved

Petra Perner Horst Perner

Activities & Achievements

Preparation of the International Conference on Mass Data Analysis of Signals and Images for Medical, Biotechnological and Chemical Applications Demonstration of the Software ProtoClass to the Participants and Industries at the conference Preparation of the Workshop Proceedings Consultation to different industries about semantic indexing of images, image analysis, and usage of the software protoclass

Events

International Conference on Mass Data Analysis of Signals and Images for Medical, Biotechnological and Chemical Applications

Publications

7.15 Contribution by MTA SZTAKI

Researchers involved

S. Fazekas and D. Chetverikov (SZTAKI), B.U.Toreyin, Y.Dedeoglu (Bilkent), T.Amiaz and N.Kiryati (TAU-Visual)

Activities & Achievements

A joint dynamic showcase demo created by SZTAKI, TAU-Visual and Bilkent. The demo involves detection and segmentation of dynamic texture in real-world videos with complex context, as well as detection of smoke and fire in real-world videos for security applications. The demo and its description were submitted and accepted for presentation at the MUSCLE demo session of CIVR 2007, Amsterdam, July 2007.

Events

Dynamic Texture Showcase will be demonstrated at CIVR 2007

7.16 Contribution by Bilkent University

Researchers involved

Enis Cetin, Behcet Ugur Toreyin, Erdem Dengel, Yigithan Dedeoglu

Activities & Achievements

Bilkent started a project for forest fire detection for Turkish Ministry of Environment and Forestry. We installed 5 cameras to hilltops of Manavgat, Antalya region, Turkey. These cameras currently monitor the surrounding hills and they transfer images to computers which have equipped with software for detecting rising smoke in the video. We successfully detected a forest fire in the region. Also, we detected several controlled test fires.

Events

Turkish Ministry of Environment and Forestry will install this detection system to several forests in Turkey

8 MUSCLE Effort Table

Person-Months report for period : May - June 2007

May - June 2007

Institute	WP1a	WP1b	WP2	WP3	WP4	WP5	WP6	WP7	Total
01 - ERCIM	2.00								2
03 - UCL	0	0	0	2	0	0	0	0	2
04 - KTH	0	0	0	0	0.2	0.8	0	0.8	1.8
05 - BILKENT	0	0.2	0.3	0	0.25	0.4	0.4	4.5	6.05
06 - VIENNA PRIP	0	0	0.31	3.19	0	0	0	2.75	6.25
07 - MTA SZTAKI	0	0	0	1	0	0	0.5	1.5	3
09 - CNR-ISTI	0	0	0	0	0.8	0	0	0	0.8
11 - TUG	0	0.07	0	2	0	0	1.97	0.15	4.19
12 - UPC	0	0	0	1	0	0	0	0.8	1.8
16 - AUTH	0	0	0	0	0	3.0357	0	0	3.035
18 - TU VIENNA IFS	0	0	0	0	2	0.75	2	0.75	5.5
22 - IBAI	0	0.07	2	0	0	0	4.61	1.99	8.67
26 - TAU-SPEECH	0	0	0	0	3.5	0	0.5	0	4
27 - TAU-VISUAL	0	0	0	1	0	0	0	2	3
31 - VTT	0	0	0	0	0	0.28	0	0	0.28
32 - INRIA Ariana	0	0.23	0	0.7	0	0	0.32	0.29	1.54
32 - INRIA Tex Mex	0	0	0	0	0	2.5	0	0	2.5
34 - LTU	0	0	0	0.975	0	0	0	0	0.975
37 - ENSEA	0	0	0	0.35	0	0	0.35	0	0.7
39 - UPS – IRIT	0	0.2	0	0.5	1.5	1	0.5	2	5.7
41 - UPMC	0	0	0	0	0	0	1	0	1
42 - NUID / UCD	0	0.1	0	0	0	0	2.1	0	2.2
Total	2.00	0.87	2.61	12.715	8.25	8.7657	14.25	17.53	66.990