MUSCLE

Network of Excellence

Multimedia Understanding through Semantics, Computation and Learning

Project no. FP6-507752

Bi-monthly progress report # 19: March – April 2007

Due date of deliverable: 15.06.2007
Actual submission date: 25.07.2007

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

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

version 1.0

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

Contents

1 Overview activities in WP1

1.1 General scientific and administrative coordination

- 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
- Preparation of the Periodic Management Report 3
- Preparation of the Periodic Activity Report 3
- Preparation of the JPA 4
- Preparation and organisation of the MUSCLE Year 3 review meeting
- Maintenance of MUSCLE website by ERCIM
- Reimbursement of MUSCLE integration expenses (fellowship, mobility support grant)
- Organisation of the Budapest meeting on 23, 24 April 2007 together with MTA Sztaki in Hungary
- IBC 2007 was contacted in order to have a MUSCLE booth during the IBC exhibition in Amsterdam in September (www.ibc.org).
- Dissemination of the MUSCLE poster
- Representation of MUSCLE to the CHORUS event on 13, 14 March 2007
- Cooperation with related European projects

2 Overview activities in WP2

2.1 Contribution by BILKENT UNIVERSITY

Researchers involved

Ugur Gudukbay, Ozgur Ulusoy, Muhammet Bastan

Activities & Achievements

We have been working on developing an automatic MPEG-7 feature (Color, Shape, and Texture) extraction tool for Bilvideo; BilVideo video database system provides full support for spatio-temporal queries that contain any combination of spatial, temporal, object appearance, external predicate, trajectory projection, and similarity-based object trajectory conditions by a rule-based system built on a knowledge-base, while utilizing an object-relational database to respond to semantic (keyword, event/activity, and category-based), color, shape, and texture queries. MPEG-7 provides a rich set of standardized tools to describe the multimedia content. Structure-based Description Schemes in MPEG-7 describe the audio-visual (AV) content from the viewpoint of its structure. They are organized around a Segment Description Scheme that represents the spatial, temporal or spatio-temporal structure of the AV content. The Segment Description Scheme can be organized into a hierarchical structure to produce a Table of Content for accessing or Index for searching the AV content. Segments can be further described on the basis of perceptual features using MPEG-7 Descriptors for color, texture, shape, motion and audio features, and semantic information using textual annotations. The output of the MPEG-7 feature extraction tool will be converted into

BilVideo knowledge-base format; this process is necessary to make feature-based querying of videos and integrate all available metadata.

Publications

A paper on this work is accepted to Fifth International Workshop on Content-Based Multimedia Indexing (CBMI'07). Paper information is as follows: Integration of structural and semantic models for multimedia metadata management. Suzanne Little, Massimo Martinelli, Ovidio Salvetti, Ugur Gudukbay, Ozgur Ulusoy, Gael de Chalendar, Gregory Grefenstette, Fifth International Workshop on Content-Based Multimedia Indexing (CBMI'07) to be held on June 25-27 2007, Bordeaux, France.

2.2 Contribution by SZTAKI

Researchers involved

Dmitry Chetverikov and Sandor Fazekas

Activities & Achievements

A study on performance evaluation of optical flow features for dynamic texture recognition has been finished, described and submitted to the Special Issue of Signal Processing: Image Communication journal on Content-Based Multimedia Indexing. The study compares texture classification performance of different sets of optical flow features for different types of optical flow, two flow confidence measures and various types of spatiotemporal division of the video data. The DynTex dynamic texture database being developed by Muscle members (Peteri, Huskies, Fazekas), has been used in the classification experiments. The DynTex database was significantly extended and updated during the reported period, which allows for more diverse and accurate testing of the dynamic texture processing methods.

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 CBIM, accepted.

2.3 Contribution by CEA

Researchers involved

Pierre-Alain Moellic

Activities & Achievements

Preparation of the ImagEVAL Workshop that will be held after the CIVR Conference (University of Amsterdam). This workshop will enable the participants to present theirs works achieved during the ImagEVAL evaluation campaign. Some participants are MUSCLE

partners. The workshop will imply proceedings. The workshop will also enable to discuss about the future of ImagEVAL and details of the next editions.

2.4 Contribution by CNR-ISTI

Researchers involved

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

Activities & Achievements

4M infrastructure has been upgraded and extended by enhancing the database, the MPEG-7 features extraction and integration. Image and audio documents types (subjects and genres) have been inserted so that multimedia documents and their metadata can be stored in the system by the selected subject (or genre). The image similarity queries have been modified so that the user can select the features for the comparison and the result is normalized. Regarding MPEG-7 features extraction, structured annotations have been included. They describe the document content ("Who", "WhatObject", "WhatAction", "Where", "When", "How" and "Why"). The user interface structure has been improved. The operations have been grouped in "drop down" menus. This unit has been strengthened to guarantee the communications and the cooperation between itself and the support for the new implemented operations. A prototype of the semantic annotation unit has been developed. It enables the user to select image regions and add annotations (inside the selected areas). The algorithm ontology unit of 4M has been designed, also starting preliminary tests. Within the fellowship program, initial concepts have been developed for implementing a hybrid approach for semantic annotation exploiting both user-driven inferencing rules and multi-level artificial neural networks. The maintenance activity of the E-Team portal has continued. In addition, a paper related to the E-Team has been accepted at CBMI 2007 in Bordeaux. In the frame of ISTI's involvement in the W3C Incubator Group on 'Multimedia Semantics', discussion continued on interoperability of multimedia. We proposed a solution to the application of an algorithm ontology to assist non-expert users in discovering and applying advanced analysis processes. To enhance interoperability the exchange or comparison of analysis procedures described using the ontology was introduced, and a preliminary evaluation was provided by including a scenario describing the application of the ontology in the field of medical image analysis.

Events

- Participation to W3C Incubator Group on Multimedia Semantics (MMSEM XG) - E-Team meeting CNR-IBaI

Publications

- Suzanne Little, Massimo Martinelli, Ovidio Salvetti, Ugur Gudukbay, Ozgur Ulusoy, Gael de Chalendar and Gregory Grefenstette. Integration of structural and semantic models for multimedia metadata management. Accepted at CMBI Conference, Bordeaux , June 25-27, 2007.

2.5 Contribution by UPC

Researchers involved

Montse Pardas, Xavier Giró, Javier Ruiz

Activities & Achievements

We have finalized the computation and translation to the HD5 format (as agreed with the other partners) of the features for the e-team *Choosing features for CBIR*. The final list of features computed is the following: *Features computed for the whole image*. All of them correspond to MPEG7 descriptors. A detailed explanation of how they are computed can be found in "B. S. Manjunath, Jens-Rainer Ohm, Vinod V. Vasudevan and Akio Yamada , Color and Texture Descriptors, IEEE Transactions on circuits and systems for video technology, vol. 11, no.6, June 2001."

- ColorStructure: Corresponds to the MPEG-7 Color Structure Descriptor
- ColorLayout: Corresponds to the MPEG-7 Color Layout Descriptor
- TextureHomogeneous: Corresponds to the MPEG-7 Homogeneous Texture Descriptor
- TextureEdgeHistogram: Corresponds to the MPEG-7 Edge Histogram Descriptor.

Features computed computed for each region. These descriptors are computed for every region of a partition of the image. In this case, for each image, a field in the hd5 format has been written for each segmentation available, that is: automatic waterfall segmentation level 1, automatic waterflall segmentation level 2, and manual segmentation when it is available. The descriptors computed are the following:

- ShapeArea: Provides information about the number of pixels associated to the region
- ShapeLocalization: Provides the following coordinates of the region: centroidX, centroidY, maxX, maxY, minX and minY
- ColorMeanVariance: Mean value and Variance in the region for each color component.
- ShapeContour: List of coordinates of the boundary points of the region
- ShapeWavelet: Contour description based on the Daubechies wavelet coefficients.

Problems

The translation to the HD5 format has taken more time than expected, but is now completed.

2.6 Contribution by TU VIENNA-PRIP

Researchers involved

Allan Hanbury, Branislav Micusik, Julian Stöttinger

Activities & Achievements

Work on the preparation of two tracks at the ImageCLEF 2007 campaign has been done. These are ImageCLEFphoto, 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 in September 2007 in Budapest has continued. A presentation on the MUSCLE evaluation events was made at the CHORUS benchmarking workshop at the INRIA Rocquencourt in March.

2.7 Contribution by ENSEA

Researchers involved

Dr. Philppe-Henri Gosselin (ENSEA), Prof. Matthieu Cord (UPMC) and Prof. John Shawe-Taylor (University of Southampton, external Partner)

Activities & Achievements

John Shawe-Taylor came to Paris (UPMC) to work with Dr. Gosselin and Prof. Cord on the current project on collaborative filtering for the Netflix contest. The aim is to propose a new approach for predicting user notes on the Netflix video catalog. Our (ENSEA & UPMC) task is to generate new features for the videos from the training set, and John Shawe-Taylor's task is to turn the binary SVM algorithm into a rank predictor. We are currently evaluating a first version of our algorithms.

2.8 Contribution by IRIT-UPS

Researchers involved

Philippe Joly, Thomas Foures

Activities & Achievements

Work for the showcase "CASEWP" (showcase n° 9): Development of a web portal dedicated to on-line evaluation of audiovisual content analysis tools. In a first step, evaluated tools are more specifically temporal segmentations of video or audio documents. The portal objectives are to provide to researchers a durable environment in which their methods can be easily evaluated and, eventually, compared to others. This kind of means already exists through various evaluation campaigns, where a precise time schedule must be followed. Time constraints are justified in order to limit the effect of over fitted parameters, but synthetic analysis of improvements on the long term is not always obvious in that context. Furthermore, the access to the content set is generally impossible after campaign deadline. The objectives are to propose a place where any identified users have an access to freely available data sets, to services providing an on-line evaluation of the submitted results, and to a public repository where they can publish these results.

Publications

CIVR 2007: "On-line Evaluation of Audiovisual Content Analysis Tools", Philippe Joly et Thomas Foures (accepted)

3 Overview activities in WP3

3.1 Contribution by TAU-VISUAL

Researchers involved

Nahum KIRYATI, Ruthy KATZ

Activities & Achievements

Creation of hardware set-up for a Visual Puzzle Solver. The purpose is to study attention and gaze control mechanisms. Unlike previous puzzle-solving studies, we use a pan-tilt-zoom camera controlled by a computer. This allows to realistically model foveation as zoom-in and zoom-out processes. In this period, we designed the color calibration and geometrical calibration schemes. We also completed initial validation of the hardware setup.

3.2 Contribution by ISTI-CNR

Researchers involved

Anna Tonazzini

Activities & Achievements

Statistical analysis of microspectroscopy signals for algae classification and phylogenetic comparison. An automatic classification scheme based on Bayesian estimation and simulated annealing is being studied, in which the algae absorption spectrum is decomposed into the absorption spectra of the pigments contained in the alga itself. The pigments, modeled as mixtures of gaussian bands, are estimated as well, based on a set of algae representative of the various phylogenetic compartments.

3.3 Contribution by TAU SPEECH

Researchers involved

Arie Yeredor

Activities & Achievements

We received reviewer reports on our paper "Blind Separation of Superimposed Shifted Images Using Parameterized Joint Diagonalization", which was submitted on January'07 to IEEE Transactions on Image Processing. In addressing the issues raised by the reviewers, we

conducted a series of additional experiments, mainly with noisy mixtures. We also added some more more illustrations of intermediate results.

3.4 Contribution by UCL

Researchers involved

Fred Stentiford, Shijie Zhang

Activities & Achievements

Work Package 3.2 – Visual Saliency

The MUSCLE sponsored **Workshop on Computational Attention** took place at Bielefeld University on 22nd March as a special session in the International Conference on Computer Vision Systems. Well over 100 attended the session at which 6 papers were presented and a further 6 posters were displayed afterwards and through the lunch period. Professor Laurent Itti from the University of Southern California gave a key note talk entitled "Combined Top-down and Bottom-up Attentional Guidance" which attracted great interest. It was clear at this workshop that many researchers are looking to attentional models to gain a better understanding of the ways humans interpret images and video. Applications included automatic image cropping and zooming, defect localization, robot vision, event detection and motion detection. The papers presented were as follows:

> *Dynamic Visual Attention: competitive versus motion priority scheme*
> *A Three-Level Computational Attention Model*
> *Pop-out and IOR in Static Scenes with Region Based Visual Attention*
> *Attention Based Auto Image Cropping*
> *Attentional Robot Localization and Mapping*
> *A Biologically Inspired Visual Selective Attention Model for Image Analysis*
> *Computational Attention for Defect Localisation*
> *Computational Attention for Event Detection*
> *Presentation Agents That Adapt to Users' Visual Interest and Follow Their Preferences*
> *An Attention Based Method For Motion Detection And Estimation*
> *Salient Visual Features to Help Close the Loop in 6D SLAM*
> *Control of Attention by Nonconscious Information: Do Intentions Play a Role?*

Joint research with INRIA on **Video Copy Detection** has produced promising results using new techniques when compared with a range of different approaches. One proposed approach makes use of the intensity profiles of frame quadrants in time. The accuracy of this method increases with the length of the clip, but the computation does not rise as rapidly. A paper entitled "Video Copy Detection: A Comparative Study" will be presented at the ACM International Conference on Image and Video Retrieval in Amsterdam on July 9th - 11th.

New attention based techniques for **Motion Estimation** have been developed that make use of static salient features to measure the detailed motion of objects and their shape. New features are generated are generated at every stage of the tracking process so that new salient regions can replace those that disappear. The approach has been shown to be more accurate in certain cases than other methods for generating motion vectors from video. A paper entitled "Motion

detection using a model of visual attention" is to be published at the International Conference on Image Processing at San Antonio on September 16th – 19th.

Events

MUSCLE sponsored [Workshop on Computational Attention and Applications](#), International Conference on Vision Systems, Bielefeld, March 21-24, 2007.

Publications

1. O Oyekoya and F W M Stentiford, "Perceptual Image Retrieval Using Eye Movements," International Journal of Computer Mathematics, vol. 84, no. 4, 2007.
2. F W M Stentiford and A Bamidele, "Attention-based colour correction," Annals of the BMVA, 2007.
3. F W M Stentiford, "Attention based Auto Image Cropping," Workshop on Computational Attention and Applications, Bielefeld, March 21-24, 2007.
4. S Zhang and F W M Stentiford, " An Attention Based Method for Motion Detection and Estimation," Workshop on Computational Attention and Applications, Bielefeld, March 21-24, 2007.
5. J Law-To, O. Buisson, L. Chen, V. Gouet-Brunet, A. Joly, N. Boujemaa, I. Laptev and F.Stentiford, "Video Copy Detection: A Comparative Study, ACM Int. Conf. on Image and Video Retrieval, Amsterdam, July 9th - 11th, 2007.
6. S Zhang and F W M Stentiford, "Motion detection using a model of visual attention," ICIP, San Antonio, September 16th – 19th, 2007.

3.5 Contribution by SZTAKI

Researchers involved

Dmitry Chetverikov, Sandor Fazekas

Activities & Achievements

Two fast, simplified versions of the program for dynamic texture detection and segmentation have been developed. The first one involves long-term learning. It is adaptive and close to real-time (5-8 frames per second.) The second one involves only short-time learning and has limited adaptivity. This version is real-time (20 fps). The fast versions provide segmentation quality similar to that of the complete, theoretically well-founded version in case when the camera motion is relatively slow and constrained.

Events

Chetverikov and Fazekas participated in the Showcase meeting in Budapest (April 23-24, 2007), where the two fast versions of the dynamic texture detector were presented. Chetverikov presented the DT segmentation program at the ACV-PRIP seminar at the Vienna University of Technology (April 20).

Publications

1. S. Fazekas and D. Chetverikov, DYNAMIC TEXTURE RECOGNITION USING OPTICAL FLOW FEATURES AND TEMPORAL PERIODICITY, MUSCLE Session of Fifth International Workshop on Content-Based Multimedia Indexing (CBIM 2007), Bordeaux, accepted.
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 CBIM, accepted.

3.6 Contribution by GET-ENST

Researchers involved

Béatrice Pesquet-Popescu, Maria Trocan, Christophe Tillier

Activities & Achievements

We have continued the investigation of sparse representations for visual content. To this end, we have proposed an algorithm for efficient quantization of subband and redundant representations, based on graph cuts. We have applied this in a first time to the contourlet decomposition (a paper was accepted to IEEE ICIP in october). Another point we worked on during this period was the three step non-linear lifting representation. The submitted paper was accepted to IEEE ICIP. It describes this non additive scheme, which allows capturing the essential features of images in the approximation subband, while reducing almost to zero the details.

Publications

1. G. Piella, B. Pesquet-Popescu, A Three-step Nonlinear Lifting Scheme for Lossless Image Compression, accepted for publication in IEEE ICIP, San Antonio, USA, Oct. 2007
2. M. Trocan, B. Pesquet-Popescu, Graph-cut rate distortion algorithm for contourlet-based image compression, accepted for publication in IEEE ICIP, San Antonio, USA, Oct. 2007

3.7 Contribution by GET-ENST

Researchers involved

Béatrice Pesquet-Popescu, Carlos Gomez

Activities & Achievements

A simple and efficient technique for face recognition was developped during this period. The main difference with previous approaches is the definition of the database. Classically, an image is exploited as a single vector, by concatenating its rows, while here we simply use all the rows as vectors during the training and the recognition stages. The new algorithm reduces the computational complexity of the classical eigenface method and also reaches a higher percentage of recognition.

Publications

C. Gomes Gascon, B. Pesquet-Popescu, "A Simple and Efficient Eigenface Method", accepted for publication at ACIVS, Delft, The Netherlands, Aug. 28-31, 2007

3.8 Contribution by TUG

Researchers involved

Martin Winter, Sandra Ober, Clemens Arth, Martina Uray, Peter Roth, Helmut Grabner, Horst Bischof

Activities & Achievements

1. In our framework combining the visual vocabulary tree with coocurrences we substituted the original 128 dimensional SIFT descriptor with the 24 dimensional PCA-SIFT descriptor. While this gives a drastic boost in performance time and reduction in storage requirements the decrease of results is negligible. The current object recognition evaluations are done on the 1000 objects of the ALOI database (original and projected onto arbitrary background). 2. We restarted working on incremental LDA learning. The combination of reconstructive and discriminative information proved to be suitable for this task, especially for large datasets as Coil100. It is possible to add new instances as well as new classes in a fast and high-quality way. 3. A major shortcoming of discriminative recognition and detection methods is their noise sensitivity, both during training and recognition. This may lead to very sensitive and brittle recognition systems focusing on irrelevant information. We developed a method that selects generative and discriminative features. In particular, we boost classical Haar-like features and use the same features to approximate a generative model (i.e. eigenimages). A modified error function for boosting ensures that only features are selected that show a good discrimination and reconstruction.

Publications

Grabner Helmut, Roth Peter M., Bischof Horst Eigenboosting: Combining Discriminative and Generative Information IEEE Conference on Computer Vision and Pattern Recognition (CVPR'07), to appear, 2007

3.9 Contribution by CNR-ISTI

Researchers involved

Umberto Barcaro, Sara Colantonio, Davide Moroni, Ovidio Salvetti

Activities & Achievements

Activity in E-Team 4: Shape modelling continued by studying a novel method for shape segmentation of heart images, based on level set shape prior techniques. The method employs mimetic criteria to get a fine initialization of the active contour. In addition, preparation of the

paper "A method for the automatic computation of the left ventricle ejection fraction" to be submitted to the Open German Russian Workshop 2007.

Moreover, work on methods for left ventricle segmentation in echocardiographic image sequences.

Publications

- Colantonio S., Moroni D., Salvetti O. – Extraction and Deployment of New Features for Cardiac Shape and Function Representation. In: The Digital Patient, I. Tollis and N. Ayache eds., Ercim News, 69, pp. 36-37, 2007 - 1. Moroni D., Colantonio S., Salvetti O., Salvetti M. - Deformable structures localization and reconstruction in 3D images. In: 2nd International Conference on Computer Vision Theory and Applications (Barcelona, 8-11 March 2007). Proceedings, pp. 215-222. A. Ranchordas and H. Araujo and J. Vitria (eds.). INSTICC, 2007

3.10 Contribution by INRIA Ariana

Researchers involved

J. Zerubia, I. H. Jermyn, A. Bhattacharya, P. Horvath, T. Peng, A. El Ghoul.

Activities & Achievements

Ting Peng, joint PhD student of INRIA Ariana and the LIAMA Institute in Beijing, has continued to work on models for road network segmentation from very high resolution (0.5m) satellite images. After extensive experimental work, she has developed models consisting of a phase field higher-order active contour prior and massive scalar field models of the image both on and off the road, possibly coupled to prior knowledge in the form of a GIS map of the the road network at an earlier date. These models produce very good segmentations of main roads at 1/8 resolution (4m), despite the large amount of image clutter that remains at this resolution, and they can be considered mature. Work has now begun on extending these models to work at full resolution, and to segmenting narrower roads. Ms Peng has developed some preliminary models for the latter by asymmetrizing the interaction in the higher-order energy term to allow it to vary with the relative orientation of the contour normal vectors. Equivalently, a term nonlinear in the inner product of the normal vectors has been added to the energy. A phase field version of this model has been formulated. The idea is to enforce road edge straightness over a larger range than the road width. Clearly this might create problems at junctions, and Ms Peng has now begun the analytical and experimental work necessary to understand the possible side-effects of the new nonlinear energy. Aymen El Ghoul, student at Sup'Com, Tunis, has begun 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. Both these cases present morphological particularities when compared to general road networks, to wit, the omnipresence of right angle junctions and straight segments in the former, and the oriented nature of the latter, which leads to small branches fusing to form larger branches. Mr. El Ghoul will extend the old and develop new models for these types of networks, and test them on a suite of images provided by the French Space Agency, CNES. His current work consists of a bibliographic study of possible modelling tools, and an assessment of the state of the art for these problems.

This will result in an internal report in June 2007. Peter Horvath, joint PhD student of INRIA Ariana and the University of Szeged, developed a multispectral image model to take full advantage of the colour infrared images used to extract tree crowns. The model is Gaussian, and includes correlations between the different spectral bands (near infrared, red, green). Experiments show that this image model, when used in conjunction with the higher-order active contour 'gas of circles' model, outperforms both greyscale higher-order models, and standard active contours that use the multispectral image model. Avik Bhattacharya, joint PhD student of INRIA Ariana and ENST, has continued his work on using road networks for retrieval from remote sensing image databases. Results on an expanded database containing a larger number of classes, using a feature set extended to treat urban and rural areas on a different footing, continue to be good. Extensive tests are now being conducted to compare the method with other existing feature sets.

In addition, scientific coordination and advising for all interns, Masters students, PhD students, and postdocs working in or with INRIA Ariana on MUSCLE topics was carried out. Coordination of MUSCLE e-team 'Shape Modelling' including budget and funding applications was achieved. 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.

Events

Urban, Paris, France (11 - 13 April 2007). CBMI, Bordeaux, France (25 - 27 June 2007). ACIVS, Delft, Netherlands (28 - 31 August 2007).

Publications

Papers have been accepted at the conferences Urban 2007 (Peng), CBMI 2007 (Bhattacharya), and ACIVS 2007 (Horvath), and to the Journal of Mathematical Imaging and Vision (Rochery) and the International Journal of Simulation Modelling (Bhattacharya).

3.11 Contribution by UPC

Researchers involved

C.C. Dorea, M. Pardas, F. Marques

Activities & Achievements

We have worked on a hierarchical analysis framework for image sequences. Region merging schemes traditionally used in the construction of partition hierarchies are extended to multiple frames using trajectory merging criteria. The merging criteria assess homogeneity among features throughout the entire sequence to recursively create partitions in the spatio-temporal domain. We propose similarity measures using long-term affine and translational motion features. Furthermore, the analysis of connectivity relations and the algorithm implementation over Trajectory Adjacency Graphs allow the generation of partition sets containing temporally consistent objects characterized by coherent motion. Lastly, we introduce the novel Trajectory Tree as a single, hierarchical representation of the partitions generated for the complete sequence. Experimental results are provided, illustrating the usefulness of the approach.

Publications

C.C. Dórea, M. Pardàs, F. Marqués, "Hierarchical partition-based representations for image sequences using trajectory merging critera", IEEE Int. Conference on Acoustics, Speech and Signal Processing, ICASSP'07, Honolulu, USA, April 2007.

3.12 Contribution by GET/ENST

Researchers involved

Y. Gousseau

Activities & Achievements

1) Matching of local, SIFT-like features (collaborators: Julien Rabin, Julie Delon). We kept on working on the automatic matching of local features. The previously developed method was extensively tested in various frameworks, and in particular in the case of multi-object matching and multi-images databases. We are currently working on the generalization of the method to color images. A paper has been writen and submitted. 2) Indexing of satellite images (coll. Jean-Francois Aujol, Bin Luo, Henri Maitre, Said Ladjal). The previously developed wavelet-based, resolution independant, indexing procedure has been applied to both classification and matching tasks on multi-resolution images. We also kept on working on the use of non-linear scale space to compute local characteristic scales on images. The resulting approach relies on a filtering of the topographic map of images (set of all level lines).

3.13 Contribution by TU Vienna-PRIP

Researchers involved

Allan Hanbury, Branislav Micusik, Lech Szumilas, Julian Stöttinger

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, particularly for bones in X-ray images. 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. The main work in this period was on the simplification of the features to improve their effectiveness and on improved difference measures between the features. These features have also been applied to encoding salient point features in medical X-ray images. Work on colour interest points has continued in cooperation with the University of Amsterdam. A number of image retrieval experiments on large image datasets have been done. In addition, the use of these interest points in object recognition has been investigated. A conference paper has been accepted for presentation at the International Conference on Image Processing (ICIP). Coordination of the E-team on "Choosing Features for CBIR and Automated Image Annotation". Work was done on the

definition of the feature exchange format as well as the consolidation of the features submitted by other E-team members.

3.14 Contribution by ARMINES-CMM

Researchers involved

Beatriz Marcotegui

Activities & Achievements

During these months we have continued collaboration with TU-Vienna-PRIP on "Matching of images using 2D colour histograms". Constitution of a personnal image database with 4 common locations: a sofa, the area in front of a house, the forest and the beach. In our database, the house images are characterised by the colour of the wall, which is close to the colour of the sofa. The image colour histogram, usually used for indexing purposes, eliminates information on the spatial distribution of colours. In order to find all the photos taken in a specific location, we perform retrieval experiments characterising images by a 2-dimensional histogram which includes information on colour adjacency. A joint paper is under preparation.

3.15 Contribution by Aristotle University of Thessaloniki

Researchers involved

I. Pitas, N. Nikolaidis, S. Asteriadis

Activities & Achievements

During the reporting period work on facial feature detection using geometrical and color/intensity information continued and concluded. This work extended previous work on this topic (eyes detection and eye centers localization) performed within MUSCLE towards mouth detection and mouth corners localization and more extensive experimentation. A face detector is applied to detect the facial region, and the edge map of this region is extracted. The distance vector field of the face is extracted in order to detect eyes and mouth. For the extraction of the distance vector field, a vector pointing to the closest edge pixel is assigned to every facial image pixel. The x and y components of these vectors are used to detect the eyes and mouth regions by comparing them with appropriate templates. Facial luminance information is used for eye center localization, after removing unwanted effects, such as light reflections. The hue channel of the lip area is used for the detection of the mouth corners. The proposed method can work efficiently both on low and high resolution images. The method has been extensively tested on the XM2VTS database (both for feature detection and localization accuracy) with very good results, that surpass results recently reported in the literature. As a final result of this work a journal paper was prepared and submitted to Pattern Recognition, where it is currently under review.

3.16 Contribution by ENSEA

Researchers involved

Dr. F. Precioso, S. Zhao (ENSEA), Prof. M. Cord (UPMC)

Activities & Achievements

Prof. Matthieu Cord and F. Precioso are currently collaborating on developing a visual semantic class learning system for video content analysis. This system is based, at a coarse level, on a robust and fast people detection using gentle AdaBoost algorithm. This detection process is refine through an accurate real-time video segmentation by smoothing spline active contours. The relevance of the segmentation allows to combine both local and global low-level features in order to provide relevant feature vectors for the learning process. To preserve the temporal coherency of video objects, we consider the spatiotemporal tube made by the segmented and characterized object along the shot sequence as one video object.

Problems

Although, this definition of a video object, with a spatiotemporal tube, is more coherent and more robust in terms of semantic class recognition, it provides new issues to be solved. Indeed, our learning system will handle the video objects using adapted kernels on set and kernels on sequence functions.

3.17 Contribution by UPS-IRIT

Researchers involved

Christine Sénac, Elie El Khoury, Nourredine Besbes

Activities & Achievements

In the context of visual speaker detection, we tried to use our audio segmentation method for shot boundary detection. We participated to the ARGOS French evaluation campaign (in the first and second phases) for the topic 'shot boundary detection' for which results for the first phase are among the best. We are waiting for the second phase results.
We started a study concerning the characters indexing in video documents. This study is based on: detecting each sequence containing a face building a model for the face of each sequence clustering the models in order to obtain in the ideal case the face of one character per cluster

3.18 Contribution by UNIS

Researchers involved

W. Christmas, F. Yan, A. Khan, B. Goswami

Activities & Achievements

Fei Yan submitted his thesis on techniques for tracking balls in sports videos.
A report is in preparation on our work on level set methods for image segmentation. We have applied them to the problem of lip tracking. Initial results on single images were promising. We plan to extend the method to image sequences. This could be either by using a 3-D space/time volume approach or by using the segmentation result from the previous frame to seed the segmentation of the current image.
We wrote a paper entitled "All Pairs Shortest Path Formulation for Multiple Object Tracking with Application to Tennis Video Analysis" for BMVC 2007.
We are investigating the use of level sets for object tracking in video sequences, with application to lip tracking.

3.19 Contribution by UvA

Researchers involved

Jasper Uijlings, Nicu Sebe

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.

Publications

1. Two-dimension Adaptive Discriminant Analysis Y. Lu, J. Yu, N. Sebe, Q. Tian, International Conference on Acoustics, Speech, and Signal Processing, Honolulu, April 2007.
2. 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.

4 Overview activities in WP4

4.1 Contribution by TAU SPEECH

Researchers involved

Arie Yeredor and David Burshtein

Activities & Achievements

1) We have revised the paper, "Support Vector Machine Training for Improved Hidden Markov Modeling," (accepted for publication in IEEE Trans. on Signal Processing in May 2007).
2) We have continued our work on the above topic by implementing the algorithm to word spotting and by evaluating the performance using standard databases.
3) We have revised the paper "Efficient Speaker Recognition Using Approximated Cross Entropy (ACE)," (accepted for publication in IEEE Transactions on Audio, Speech and Language Processing in May 2007).
4) We continued our activities on single-channel separation of audio sources, mainly by studying and partly implementing (for further integration) two competing approaches: one by Bach and Jordan, based on spectrogram segmentation, and the other by Virtanen, combining non-negative matrix factorization with sparseness and temporal continuity objectives.

Publications

The journal papers that were noted above in the activities section will be indicated in the May-June report.

4.2 Contribution by TU Vienna-IFS

Researchers involved

Jakob Frank, Thomas Lidy, Rudolf Mayer, Andreas Rauber

Activities & Achievements

TU Vienna-IFS developed a Web Service for Audio Feature Extraction, which enables researchers and other interested people to extract audio features from their own audio files via a comfortable web service, where the processing is run entirely on our servers. Furthermore

we started the development of a web service for the training of Self-Organising Music Maps. The goal of this is to provide an integrated service that enables the submission of mp3 files and the automatic feature extraction and creation of a Music Map interface for interaction with music collections. We investigated the combination of symbolic and audio music features in a music genre classification system and submitted a paper on "Improving Genre Classification By Combination Of Audio And Symbolic Descriptors Using A Transcription System" to the ISMIR 2007 conference. Furthermore, we researched on Chord-Detection methods from polyphonic audio, and wrote a paper on this topic. A joint paper was written with TU Vienna-IFS and ISTI-CNR on the application of perceptual features for improved transient restoration of audio, which was also submitted to the ISMIR 2007 conference. TU Vienna-IFS is organizing the International Conference on Music Information Retrieval 2007 which includes several activities. In April, TU Vienna-IFS organized the MIREX 2007 European Planning meeting, discussing topics and tasks for the forthcoming Music Information Retrieval Evaluation eXchange benchmarking event.

Events

MIREX 2007 European Planning meeting, Vienna, Austria, April 21 - 23, 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. Veronika Zenz and Andreas Rauber. "Automatic Chord Detection Incorporating Beat And Key Detection". Submitted.

4.3 Contribution by CNRS LLACAN

Researchers involved

Fathi Debili, Zied Ben Tahar, Emna Souissi

Activities & Achievements

How could we produce annotated texts massively with optimal efficiency, reproducibility and cost? Instead of correcting the output of the automatic analysis with dedicated tools, as suggested currently, we found it more advisable to use interactive tools for analysis, where manual editing is fed in real time into automatic analysis. We address the issue of evaluating these tools, along with their performance in terms of linguistic ergonomy, and propose a metric for calculating the cost of editing as a number of keystrokes and mouse clicks. By way of a simple protocol addressing Arabic vowelization, tagging and lemmatization, we discover that, surprisingly, the best interactive performance of a system is not always correlated to its best automatic performance. In other words, the most performing automatic linguistic behavior of a system does not always yield the best interactive behavior, when manual editing is involved.

Publications

Final version of: "Analyse automatique vs analyse interactive : un cercle vertueux pour la voyellation, l'étiquetage et la lemmatisation de l'arabe", TALN 2007, Toulouse, 5-8 Juin 2007.

4.4 Contribution by CNR-ISTI

Researchers involved

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

Activities & Achievements

Preliminary subjective and objective tests of ARIA algorithm Winamp plug-in (using ISMIR mp3 genre collection) for transients enhancement of compressed music and developpement of an automatic adaptation method. Final test and integration of a software library for real-time extraction of pitch and RMS of musical signals into a multimedia authoring environment. Preparation of papers to be submitted to conferences. Progress in the activities of MODEM EU project, a web-based environment for exchanging multi-track audio/musical signals development.

Problems

Delay in MUSCLE activities due to involvement of ISTI Audio team in other projects.

Publications

Papers submitted to DAFX 07, ISMIR conferences.

4.5 Contribution by Aristotle University of Thessaloniki

Researchers involved

C. Kotropoulos, D. Ververidis

Activities & Achievements

The relationship between the variance of the prediction error committed by the Bayes classifier and the mean prediction error was established by experiments in emotional speech classification within a cross-validation framework in a previous work. Further research has justified theoretically the validity of the aforementioned relationship. We have proved that by treating the new estimate of the variance of the prediction error as a random variable, it exhibits a much smaller variance than the usual estimate obtained by cross-validation even for a small number of repetitions.

4.6 Contribution by IRIT-UPS

Researchers involved

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

Activities & Achievements

A collaboration with UPC (Barcelona, Spain) started in April. Andrey Temko, a PhD student at UPC, joined IRIT for a period of 3 months. The goal of this collaboration is to apply the sequence kernels developed at IRIT to the problem of acoustic events detection. A new visit of the Speech/nonspeech, Music/nonmusic and Song/nonSong detections is proposed. Two communications are proposed and accepted (GRETSI07, EUSIPCO07)

Problems

The accuracy of current acoustic events detection systems is very low (less than 30%). The major problem is to detect the events that are overlapped with speech.

Publications

Singing Voice Characterization For Audio Indexing, Hélène Lachambre, Régine André-Obrecht, and Julien Pinquier, EUSIPCO '07, September 2007 (accepted)

5 Overview activities in WP5

5.1 Contribution by INRIA-TEXMEX

Researchers involved

Patrick Gros, Xavier Naturel, Guillaume Gravier, Ewa Kijak

Activities & Achievements

1. Following the thesis of Manolis Delakis, we worked on a journal paper of his work. The paper was submitted to the "Computer Vision and Image Unnderstanding" journal. Revisions were asked by the reviewers and a corrected version was submitted.
2. Patrick Gros is co-editor of the book planned by WP5 on multimedia analysis and interfaces. A meeting in March in Athens for a showcase was the occasion to advance on several issues of this book.
3. We are involved in three showcases of WP5. A) Movie Summarization. A meeting took place in March in Athens to refine the scenario of the showcase, and the specification of the module we should develop was done. B) ACADI. TexMex defined and provided the video corpus. We have run our face detector on the whole corpus and the results were used as input for costume detection. We have supplied a server and client library used in the demonstrator that offers many functionalities for video decoding and access. A teleconference was held on March 30th, and a meeting took place in Rennes, on April 17th. C) AV ASR. During the stay of Guillaume Gravier at NTUA (April 24-27), we had discussions on how to enhance the

AVSR showcase using improved technology (e.g. variance compensation, landmark driven decoding). This visit was also the opportunity to exchange some data in order to work with a common database and protocol, thus enabling a closer collaboration of the two groups.

Events

- March 15th. Movie summerization meeting, Athens.
- April 17th. ACADI meeting, Rennes.
- April 24-27th. Visit of G. Gravier to NTUA.

Publications

Xavier Naturel, Patrick Gros. Detecting Repeats for Video Structuring. Multimedia Tools and Application, To appear.

5.2 Contribution by TU Vienna-IFS

Researchers involved

Robert Neumayer, Andreas Rauber

Activities & Achievements

TU Vienna-IFS continued research on the combination of textual and audio features for Music Information Retrieval problems. We extended 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. Our paper on "Multi-modal music information retrieval - visualisation and evaluation of clusterings by both audio and lyrics" was accepted at the 8th Conference on Recherche d'Information Assistée par Ordinateur (RIAO'07).

Events

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

5.3 Contribution by Technion - MM

Researchers involved

Ron Kimmel and Ehud Rivlin

Activities & Achievements

Most of our work in this period was devoted for human body parts tracking in 3D with pre-learned motion models using multiple cameras. We collaborated closely on this with UPC. Results of this work were presented in the following conferences: 2nd International Conference on Computer Vision Theory and Applications (VISAPP 2007) in Barcelona, The 5th International Conference on Computer Vision Systems (ICVS07) in Bielefeld, Germany. In this work we present an approach for human body parts tracking in 3D. Human body pose estimation and tracking is a challenging task for several reasons. First, the large dimensionality of the human 3D model complicates the examination of the entire subject and makes it harder to detect each body part separately. Secondly, the significantly different appearance of different people that stems from various clothing styles and illumination variations, adds to the already great variety of images of different individuals. Finally, the most challenging difficulty that has to be solved in order to achieve satisfactory results of pose understanding is overcoming the ambiguity in solving for the parameters. We present an approach for human body parts tracking in 3D with prelearned motion models using multiple cameras. We proposed Gaussian Process Annealing Particle Filter GPAPF for tracking in order to reduce the dimensionality of the problem and to increase the tracker's stability and robustness in tracking objects in high dimensional space. The annealed particle filter has a good performance when applied on videos with a high frame rate, but performance drops when the frame rate is lower. We show that our approach provides good results even for the low frame rate (30fps and lower). We have also shown that using Gaussian Process Dynamic Model GPDM can increase the ability to recover from temporal target loss. We have also presented a method of evaluation the possibility of self occlusion and we have suggested a way to adjust the weighed function for such cases, in order to be able to produce more accurate evaluation of a pose.The main problem of our method is that the learning and tracking are done for a specific action. The ability of the tracker to use a latent space in order to track a different motion type, has not been shown yet. A possible approach is to construct a common latent space for the poses from different actions. The difficulty with such approach may be the presence of a large number of gaps between the consecutive poses. In the future we plan to extend this approach in order to be able to track different activities, using same learned data. Other challenging task is to track two or more people simultaneously. The main problem here is that in this case there is high possibility of occlusion. Furthermore, while for a single person each body part can be seen by at least one camera that is not the case for the crowded scenes.

Publications

2nd International Conference on Computer Vision Theory and Applications (VISAPP 2007) in Barcelona, The 5th International Conference on Computer Vision Systems (ICVS07) in Bielefeld, Germany.

5.4 Contribution by Aristotle University of Thessaloniki

Researchers involved

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

Activities & Achievements

AUTH MUSCLE Movie Database Release 1.1 is a test-bed proposal for deeper annotation. The database contains scenes from 6 movies. In total, 25 dialogue and 17 non-dialogue scenes are currently available, where dialogues last from 20 sec to 123 sec. The audio track of these scenes is digitized in PCM, at a sampling rate of 48 kHz and each sample is quantized in 16 bit two-channel. The video data are encoded using the DivX 6.1.1 encoder and are stored in .avi format. Frame dimensions are either 720x576 or 660x384 and the frame rate is 25 fps. Depending on the original recording aspect ratio, black stripes of variable dimensions exist on the top and bottom part of the video. The total duration of the 42 scenes is 34 min and 43 sec. The database has been shared with WP5 MUSCLE members. A module for speaker clustering based on clustering ensembles has been developed. Clustering ensembles combine the results of "dummy" clustering algorithms, so that the final clustering solution reveals the natural groupings of the data. In our implementation, three "dummy" hierarchical clustering algorithms are used, namely, the weighted linkage algorithm, the average linkage algorithm, and the Ward's method. In total, 30 clustering solutions are produced using the aforementioned algorithms. The clustering algorithm that is used in each run is randomly selected. Afterwards, the 30 clusterings are combined and a co-association matrix is built. The final clustering is performed on the co-association matrix, that reveals the natural clusters of the data. During the reporting period an algorithm that clusters face images found in video sequences was investigated. The algorithm is based on a novel method for creating a dissimilarity matrix using SIFT image features. This dissimilarity matrix is used as an input in a hierarchical average linkage clustering algorithm, which yields the clustering result. Applications of the method include dialogue detection (which is the actual application where this method will be incorporated within MUSCLE along with speech-related information), video indexing and content analysis and as a pre-processing step for face recognition. This work is still in progress.

Events

C. Kotropoulos, V. Moschou and P. Antonopoulos attended the WP5-eteam and showcase workshop in Athens on March 15, 2007.

5.5 Contribution by IRIT-UPS

Researchers involved

Julien Pinquier, Eduardo Sanchez-Soto, Frederic Gianni, Christine Sénac, Elie El Khoury.

Activities & Achievements

I) During this period, several important points of the showcase n°8 (ACADI for Automatic Character in Audiovisual Document Indexing) were approached: - tool adaptation, - common exchange format, - video management and decoding. This work, between SAMoVA and Texmex teams, was completed through 1 phone meetings (March 30th, to discuss about the exchange format and the first results of each tool) and 1 meeting at the IRISA (April 17th, to choose Texmex video manager called "DIVA" and to prepare tool integration).

II) The goal of the showcase « Real-Time Audio-Visual Automatic Speech Recognition Demonstrator is to build and proof a practically deployable Audio-Visual Speech recognition prototype. Our work in this show case, where the Technical University of Crete is the leader, was to develop the programs and scripts (employing the HTK tool kit) to perform the back-end process, principally the recognition stages. At the same time we have developed a basic GUI (graphical user interface) using Perl/Tk to interact with this part of the system.

III) We continued our work on unsupervised optimal stream weight computation. We proposed a first extension of our previous work on classification to handle the problem of detection/segmentation. We applied it to an audio-visual speech recognition task. The first results were submitted to ICSPC'2007.

Publications

Julien Pinquier, ACADI: Automatic Character (in Audiovisual Document) Indexing. Application in: Large-Scale Semantic Access to Content (Text, Image, Video and Sound) – RIAO2007, Pittsburgh, USA, 30/05/2007-01/06/2007. Frédéric GIANNI, Julien PINQUIER and Ewa Kijak, ACADI Showcase - Automatic Character Indexing in Audiovisual Document. In: ACM International Conference on Image and Video Retrieval - CIVR 2007, Amsterdam, The Netherlands, July 9-11 2007.

6 Overview activities in WP6

6.1 Contribution by TU Vienna-IFS

Researchers involved

Rudolf Mayer, Robert Neumayer, Andreas Rauber

Activities & Achievements

Participation in the e-team meeting on "Active Semi-Supervised Learning" at INRIA in Rocquencourt (France), in March 13 - 15, 2007, discussing the Roadmap of the e-Team, how to keep the user's interest during interaction with active learning, multiple contexts of use and the difficulty of benchmarking. Rudolf Mayer gave a talk about current work on the user-based "correction" of Self-organizing feature maps, currently in a proof-of-concept stage. Several E-Team Collaborations started, one between INRIA-IMEDIA and TU Vienna-IFS will continue investigating correcting and updating SOMs; an e-Team exchange is planned for May. A paper on a novel visualisation method of component planes for self-organising maps ("The metro visualisation") was acceped at the International Joint Conference on Neural Networks (IJCNN'07). A paper on "Visualising Class Distribution on Self-Organising Maps" was accepted for publication at the International Conference on Artificial Neural Networks (ICANN'07).

Events

E-team meeting "Active Semi-Supervised Learning" at INRIA in Rocquencourt (France), March 13 - 15, 2007

Publications

- Robert Neumayer, Rudolf Mayer, Georg Pölzlbauer, and Andreas Rauber. The metro visualisation of component planes for self-organising maps (accepted for publication). In Proceedings of the International Joint Conference on Neural Networks (IJCNN'07), Orlando, FL, USA, August 12 - 17 2007. IEEE Computer Society.
- Rudolf Mayer, Taha Abdel Aziz, and Andreas Rauber. Visualising Class Distribution on Self-Organising Maps (accepted for publication). In Proceedings of the International Conference on Artificial Neural Networks (ICANN'07), Porto, Portugal, September 9 - 13 2007. Springer Verlag.

6.2 Contribution by TUG

Researchers involved

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

Activities & Achievements

1. Based on the proposed probabilistic tracking algorithm developed in the past different methods where implemented and tested to increase the robustness of the tracker. Segmentation of objects based on tracking results is made to generate accurate real world positions of athletes, which are further evaluated by sport scientists.
2. We finished the high performance C++ implementations of our developed clustering algorithms and are currently experimenting with clustering large scale datasets, specially from visual object recognition domain.

Publications

Mauthner Thomas, Bischof Horst A Robust Multiple Object Tracking for Sport Applications Performance Evaluation for Computer Vision, 31st AAPR/OAGM Workshop 2007, p.81-89, Austrian Association for Pattern Recognition, Oesterreichische Computer Gesellschaft, 2007-May.

6.3 Contribution by INRIA Ariana

Researchers involved

J. Zerubia, I. H. Jermyn

Activities & Achievements

The higher-order active contour (HOAC) model for a 'gas of circles', developed during the PhD of Peter Horvath, joint PhD student in INRIA-Ariana and the University of Szeged, has been further improved by reformulating it as a nonlocal phase field model, in the same way as was done for the original higher-order active contour road network models. The translation of the parameters from contour to phase field works well, and enables the preservation of the inflection point constraint that allows all but one of the prior model parameters to be fixed.

The results obtained in the tree crown extraction problem are of the same high quality as those obtained with the contour model, but the computation time has been reduced by two orders of magnitude.

Events

CAIP, Vienna Austraia (27 - 29 August 2007). EUSIPCO, Poznan, Poland (3 - 7 September 2007).

Publications

Papers have been accepted at the conferences CAIP 2007 (Horvath) and EUSIPCO 2007 (Horvath).

6.4 Contribution by CNR-ISTI

Researchers involved

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

Activities & Achievements

Within the activities of E-team 15, the organization of the invited session at the 11th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems has been finalized, with 5 accepted contributions (3 from within Muscle). Research on Bayesian blind separation techniques for astrophysical imaging has produced an almost complete definition of a specific source and data statistical model. Study and development of methods for integrating inferential and computational reasoning for supporting clinical decision making. The computational reasoning module is able to trigger both artificial intelligence algorithms (namely support vector machine) and image feature extraction algorithms. The architecture has been instantiated to fulfil the requirements of the cardiological field. In addition, identification of decision support system architecture for image-based clinical diagnosis.

Publications

- V. Schetinin, V. V. Zharkova, A. Brazhnikov, S. Zharkov, E. Salerno, L. Bedini, E. E. Kuruoglu, A. Tonazzini, D. Zazula, B. Cigale, H. Yoshida, "Advanced Feature Recognition and Classification Using Artificial Intelligence Paradigms", in V. Zharkova and L. C. Jain, Eds., Artificial Intelligence in Recognition and Classification of Astrophysical and Medical Images, Springer-Verlag, SCI Series, Vol. 46, Berlin, Heidelberg, 2007, pp. 151-338
- Moroni D., Colantonio S., Salvetti O., Salvetti M. - Deformable structures localization and reconstruction in 3D images. In: 2nd International Conference on Computer Vision Theory and Applications (Barcelona, 8-11 March 2007), pp. 215-222. A. Ranchordas and H. Araujo and J. Vitria (eds.). INSTICC, 2007

6.5 Contribution by UCD

Researchers involved

Pádraig Cunningham, Ken Bryan, Derek Greene

Activities & Achievements

Work on the Muscle Book on Machine Learning and Multimedia continues. The current plan is to have this book completed by the end of July. A paper entitled: "An Ensemble Approach to Identifying Informative Constraints for Semi-Supervised Clustering" by Derek Greene, Pádraig Cunningham has been submitted to ECML. We have been active in research on one-class classification, specifically on dimension reduction for one class classification and on the use of one-class classification methods for speaker identification.

Events

Pádraig Cunningham attended the Muscle meeting in Budapest in April.

Publications

The following paper has been accepted for publication at the 5th IEEE International Workshop on Content-Based Multimedia Indexing: "An assessment of alternative strategies for constructing EMD-based kernel functions for use in an SVM for image classification" by Anton Zamolotskikh and Pádraig Cunningham

6.6 Contribution by Aristotle University of Thessaloniki

Researchers involved

M. Sedaaghi (MUSCLE fellow) C. Kotropoulos, D. Ververidis

Activities & Achievements

Self-adaptive genetic algorithms are employed to search for the worst performing features with respect to the probability of correct classification achieved by the Bayes classifier in a first stage. That is, a genetic algorithm-based implementation of backward feature selection is proposed. 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. The sequential application of both stages is demonstrated to improve speech emotion recognition on the Danish Emotional Speech database. A chapter on theoretical analysis supporting the development of wrappers with low computational cost has been contributed to WP6 deliverable on Supervised Learning.

6.7 Contribution by ENSEA

Researchers involved

Prof. S. Philipp-Foliguet, Dr. F. Precioso, Dr. M. Jordan, J. Lebrun, M. Alcoverro (All from ENSEA), L. Najman and M. Couprie (Marne-la-Vallée University, external collaborator)

Activities & Achievements

We have had several work meetings with L. Najman and M. Couprie from A2SI team of Gaspard Monge CNRS Lab (Marne-la-Vallée), who are specialists of discrete geometry and have developed several algorithms for watershed. With them, we have studied and implemented a new method to segment 3D objects, based on a filtering of the surface followed by a watershed segmentation using the local curvature. This segmentation into surfacic regions will allow us to build signatures which will be used to compare and to classify artwork objects coming from the Louvre museum. This work directly concerns to improve our methods, involved in MUSCLE project, regarding 3D data indexing and retrieving.

6.8 Contribution by IRIT-UPS

Researchers involved

Khalid Daoudi, Jérôme Louradour

Activities & Achievements

A software for SVM classification with sequence kernels has been developed. This software allows the use and comparison of several state-of-the-art sequence kernels: Fisher kernel, probability-product kernels, supervectors GMM kernels, GLDS kernel and the FSMS kernels developed at IRIT. This software is the result of the research effort made in the E-team "Kernel methods in dynamic classification problems" and is made available for all MUSCLE partners. A documentation is available but only in French for the moment.

Events

On April 10th, Khalid Daoudi made a talk at Google (Mountain-View, CA, USA) entitled "On Sequence Kernels for SVM classification of sets of vectors". This talk covered the research made in E-team 13. A video of the talk can be found on http://video.google.com/videoplay?docid=6518731752586282105&q=Khalid+Daoudi

Publications

The journal paper submitted to IEEE Transactions on Audio, Speech and Language Processing entitled "Feature Space Mahalanobis Sequence Kernels: Application to SVM Speaker Verification" by J. Louradour, K. Daoudi and F. Bach has been "accepted for publication with mandatory minor revisions" in April.

6.9 Contribution by UTIA

Researchers involved

M. Haindl, S. Mikes

Activities & Achievements

A novel unsupervised multi-spectral multiple-segmenter texture segmentation method with unknown number of classes was developed and extensively tested. This unsupervised segmenter is based on a combination of several unsupervised segmentation results, each in different resolution, using the sum rule. Multi-spectral texture mosaics are locally represented by four causal multi-spectral random field models recursively evaluated for each pixel.

Publications

MP-576

6.10 Contribution by IBaI

Researchers involved

Petra Perner, Horst Perner

Activities & Achievements

We have been working on a conceptual approach for novelty detection and handling based on case-based reasoning. Different scenarios for novelty detection have been worked out. Methods for handling known situation and detect unknown stituation were developed as well as methods for handling unknown situation and making them immediately available for reasoning. Besides we been working on a book for Case-Based REasoning on Signals and Images that will be published by Springer in the Series Computational Intelligence.

Events

Industrial Conference on Data Mining ICDM 2007 www.data-mining-forum.de

Publications

Concepts for Novelty Detection and Handling based on a Case-Based Reasoning Process Scheme P. Perner Springer Verlag to appear

6.11 Contribution by UNIS

Researchers involved

I. Kolonias, W. Christmas, J. Kittler

Activities & Achievements

We attended the International Conference on Visual Systems in Bielefeld, Germany to present the paper listed below. This paper described some of the architectural aspects of our work on automatic annotation of tennis videos. The layered memory architecture was inspired by the human visual system, with its sensory, short term, working and long-term memory. We are continuing the work described in this paper in order to publish an expanded version in a journal.

We have worked on extending the application of top-down consistency validation enforcement for events detected in tennis video sequences. More specifically, we implemented the grammar that corresponds to the award of games from individual points scored in the tennis rules, and examined how enforcing this grammar on the tennis video evolution affects the overall system performance. The process followed and the results obtained have been included in a paper entitled "Improving the accuracy of automatic tennis video annotation by high level grammar" which has been accepted for presentation at the 2007 International Conference on Image Analysis and Processing in Modena, Italy on September.

Publications

Ilias Kolonias, William Christmas, and Josef Kittler. A layered active memory architecture for cognitive vision systems. In Proc. ICVS 2007, March 2007.

7 Overview activities in WP7

7.1 Contribution by MTA SZTAKI

Researchers involved

László Havasi Levente Kovács István Petrás Tamás Szirányi Zoltán Szlávik

Activities & Achievements

Visual surveillance and activity analysis is an active research field of computer vision. As a result, there are several different algo-rithms 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. 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: a) the client's web interface, b) the server (possibly but not necessarily including the web server) c) the controller and d) the communication inter-face embedded into the user module. Each component operates autonomously communication through RPC requests over TCP/IP.

Problems

Visual surveillance and activity analysis has attained great interest in the field of computer vision research. Several algorithm libraries are available on-line (open-source or proprietary),

how-ever their integration into a complex system is hindered by the inhomogeneity of the implementation language, format, process-ing speed, etc. The aim of this work is to produce a flexible, transparent system for activity analysis. The system provides a transparent interface to heterogeneous modules with different input-output requirements. The setup is hierarchical thus helping the scalability of the whole framework. The actual implementation integrates diverse algorithms forming a test-bed for unusual activ-ity detection. Various complex surveillance related algorithms, such as human and body action, tracking and motion activity algorithms are integrated into one system. What is unusual? An event that occurs rarely. We interpret the term unusual in statistical sense. Intelligent visual surveillance is an increasingly important part of computer vision research. One of the most important goals of visual surveillance systems is to analyze the activity of the ob-served objects in order to detect anomalies, predict future behaviours, or predict potential unusual events before they occur. There have been a lot of approaches to model the activity of dynamic scenes. Analysis of motion patterns is an effective approach for learning the observed activity. For the most of the time, objects in the scene do not move randomly. They usually follow well-defined motion patterns. Knowledge of usual motion patterns can be used to detect anomalous motion patterns of objects. Current systems mainly base their analysis of motion patterns on a predefined classification of tracked data or of optical flow patterns. The module consists of three processing sub-modules. First, the input image sequence is filtered to remove noise and enhance significant edges. Then the EMD (elementary motion detection module estimates the local motion vectors in every pixel which are fed into the GMM (Gaussian mixture model) sub-module. The GMM module learns the usual motion patterns and evaluates whether the input pixel values are from the learned sets or not. Finally, the Decision module evaluates the output of GMM sub-module and produces an alarm signal if an unusual motion pattern was detected.

Events

CIVR: ACM International Conference on Image and Video Retrieval, July 9-11 2007

Publications

CIVR: ACM International Conference on Image and Video Retrieval, July 9-11 2007 Losteiner, D. et al., "Processing images of a moving camera", Sixth Conf. of Hungarian Association for Image Processing and Pattern Recognition, Debrecen, pp. 283-291, 2007

Researchers involved

Dmitry Chetverikov and Sandor Fazekas

Activities & Achievements

The original version of the program for dynamic texture detection and segmentation developed jointly by SZTAKI and TAU-Visual is relatively slow (15-20 secons per frame). Two fast, simplified versions of the program have been developed for the MUSCLE Dynamic Texture Showcase joint with Bilkent. The first version is adaptive and close to real-time (5-8 frames per second), the second one has limited adaptivity but is real-time (20 fps). The fast versions provide segmentation quality similar to that of the original version in case when the camera motion is relatively slow and constrained.

Events

Chetverikov and Fazekas participated in the Showcase meeting in Budapest (April 23-24, 2007), where the two fast versions of the dynamic texture detector were presented. Chetverikov presented the DT segmentation program at the ACV-PRIP seminar at the Vienna University of Technology (April 20).

Publications

A joint showcase paper (Bilkent, SZTAKI, TAU-Visual) has been accepted for CIVR 2007 (Amsterdam): Toreyin, Dedeoglu, Cetin, Fazekas, Chetverikov, Amiaz, Kiryati: Dynamic Texture Detection, Segmentation and Analysis

7.3 Contribution by GET-ENST

Researchers involved

Béatrice Pesquet6Popescu, Maria Trocan, Christophe Tillier

Activities & Achievements

We continued the work in collaboration with Bilkent University on copy detection and flame detection. On each of these subjects, a couple of master students have started the implementation of the proposed methods: - For flame detection, we want to exploit the motion vector fields in a spatio-temporal representation. Different methods for motion estimation will be explored and also different coherence parameters between motion vector fields - For copy detection, we also try to exploit the temporal information, which is more "stable" to format changes than the texture in video.

7.4 Contribution by TU Vienna-IFS

Researchers involved

Jakob Frank, Thomas Lidy, Andreas Rauber

Activities & Achievements

TU Vienna-IFS coordinated the presence of MUSCLE with a booth at the CeBIT fair in Hannover (Germany) from March 15 - 21, which was a major event and successful for dissemination of MUSCLE research results. Several showcases were presented with live demos and a projection wall as well as leaflets were available for presentation of all MUSCLE research activities. A number of interesting contacts and ideas were collected during CeBIT. For the Content Analysis Showcase we did a segmentation of the audio tracks from the videos in the showcase database. From the segmented audio tracks we extracted audio features and visualized the distribution of segments on a Self-Organizing Map. Progress was made on the MUSCLE Showcase on "Shaping 3-dimensional Multimedia Environments" in collaboration with EC3. A book chapter was written about the 3D Multimedia Environment and submitted

to "Computational Intelligence in Multimedia Processing: Recent Advances", to be published by Series "Studies in Computational Intelligence", Springer Verlag. A MUSCLE Showcase meeting was held in Budapest, Hungary on April 23-24, 2007, where progress on both showcases was presented.

Events

CeBIT in Hannover March 15 - 21, 2007 MUSCLE Showcase Progress Meeting, Budapest, Hungary, April 23-24, 2007

Publications

Ronald Genswaider, Helmut Berger, Michael Dittenbach, Andreas Pesenhofer, Thomas Lidy, Andreas Rauber, and Dieter Merkl. "Shaping 3-dimensional Multimedia Environments". In Computational Intelligence in Multimedia Processing: Recent Advances, Springer Series on Studies in Computational Intelligence, Springer Verlag. Accepted for publication.

7.5 Contribution by INRIA Ariana

Researchers involved

J. Zerubia, I. H. Jermyn.

Activities & Achievements

Ian Jermyn gave an invited talk at the French Space Agency CNES Orfeo Symposium, dedicated to the Orfeo Accompaniment Programme for the new Pleiades very high resolution satellites. He visited Professor Anuj Srivastava of the Department of Statistics at Florida State University as part of a collaboration within INRIA Associated Team 'Shapes', and gave an invited talk as part of the 'Shape Day' symposium organized by the Department of Statistics. He attended Urban 2007 to present work performed within Muscle. Josiane Zerubia and Ian Jermyn presented Muscle work to representatives of the French Space Agency, CNES, with whom INRIA Ariana have a contract to continue Muscle related work, and to transfer it to the CNES developed ORFEO Toolbox. In collaboration with Alcatel Alenia Space (now Thales Alenia Space), INRIA Ariana applied successfully for a PhD grant to help support Aymen El Ghoul, who will continue as a PhD student working on the CNES contract.

Events

CNES ORFEO Symposium, Paris, France (17 January 2007).
'Shape Day' Symposium, Tallahassee, Florida, USA (6 April 2007)

7.6 Contribution by UPC

Researchers involved

Montse Pardas, Cristian Canton

Activities & Achievements

As a contribution to the Showcase "Unusuat event detection" we have worked on the adaptation of body and motion analysis tools to be used in this framework. The system is based on three modules:

- 2D Body model. The inputs to this module are the foreground regions previously extracted by the background /foreground separation. In order to extract a set of features describing the body of a person that performs an action, a geometrical configuration of human body must be considered. Since the aim of our research is to increase robustness of gesture classification by embedding human body configuration information in our data analysis loop while keeping real-time performance, a simple elliptic model of human body has been adopted.
- Motion modeling. In order to achieve a simple and efficient motion representation, we use the concept of Motion History Image (MHI) and Motion Energy Image (MEI) introduced by Bobick.
- Unusual event detection. Data produced by the body and motion analysis modules is processed in order to extract a vector of features for classification. Informative features derived from the analyzed data (MHI and MEI) are required. Statistical moments invariant to scaling, translation, rotation and affine mappings have been used. For each data set (MHI and MEI), 5 invariant moment-based features are computed. Information from the elliptic body model can be used to generate additional features. We use two features describing the relative amount of motion pixels located in the upper and lower body part, defined through the elliptic model. Thus, we construct a 12-dimensional feature vector. For each scenario, this feature vector is trained for the usual events (people walking and people standing for instance) using a mixture of Gaussians probability model. The detection of unusual events is based on a classification of each feature vector as belonging to this model or not.

7.7 Contribution by Technion - MM

Researchers involved

Ehud Rivlin, Ron Kimmel

Activities & Achievements

Most of our work in this period was devoted for human body parts tracking in 3D with pre-learned motion models using multiple cameras. We collaborated closely on this with UPC. Results of this work were presented in the following conferences: 2nd International Conference on Computer Vision Theory and Applications (VISAPP 2007) in Barcelona, The 5th International Conference on Computer Vision Systems (ICVS07) in Bielefeld, Germany. In this work we present an approach for human body parts tracking in 3D. Human body pose estimation and tracking is a challenging task for several reasons. First, the large dimensionality of the human 3D model complicates the examination of the entire subject and makes it harder to detect each body part separately. Secondly, the significantly different appearance of different people that stems from various clothing styles and illumination variations, adds to the already great variety of images of different individuals. Finally, the

most challenging difficulty that has to be solved in order to achieve satisfactory results of pose understanding is overcoming the ambiguity in solving for the parameters. We present an approach for human body parts tracking in 3D with prelearned motion models using multiple cameras. We proposed Gaussian Process Annealing Particle Filter GPAPF for tracking in order to reduce the dimensionality of the problem and to increase the tracker's stability and robustness in tracking objects in high dimensional space. The annealed particle filter has a good performance when applied on videos with a high frame rate, but performance drops when the frame rate is lower. We show that our approach provides good results even for the low frame rate (30fps and lower). We have also shown that using Gaussian Process Dynamic Model GPDM can increase the ability to recover from temporal target loss. We have also presented a method of evaluation the possibility of self occlusion and we have suggested a way to adjust the weighed function for such cases, in order to be able to produce more accurate evaluation of a pose. The main problem of our method is that the learning and tracking are done for a specific action. The ability of the tracker to use a latent space in order to track a different motion type, has not been shown yet. A possible approach is to construct a common latent space for the poses from different actions. The difficulty with such approach may be the presence of a large number of gaps between the consecutive poses. In the future we plan to extend this approach in order to be able to track different activities, using same learned data. Other challenging task is to track two or more people simultaneously. The main problem here is that in this case there is high possibility of occlusion. Furthermore, while for a single person each body part can be seen by at least one camera that is not the case for the crowded scenes.

Publications

Results of this work were presented in the following conferences: 2nd International Conference on Computer Vision Theory and Applications (VISAPP 2007) in Barcelona, The 5th International Conference on Computer Vision Systems (ICVS07) in Bielefeld, Germany.

7.8 Contribution by TU VIENNA-PRIP

Researchers involved

Allan Hanbury, Branislav Micusik, Julian Stöttinger

Activities & Achievements

Work has advanced on the two showcases in which we are involved.
For the evaluation showcase, many basic decisions on the running of the image retrieval showcase were made. In addition, we put together the image dataset to be used in the showcase. This consists of the IAPR-TC12 dataset (used in ImageCLEF) along with some "poorly" annotated images. This simulates common annotation practice of users, and makes the event more challenging as the systems cannot rely on good text annotations for all images relevant to a query.
For the object recognition showcase, progress was made on the conversion of the code for colour interest point detection from MATLAB to C++. In addition, initial object recognition experiments using these interest points were done. A preliminary web interface to the program was also developed. Allan Hanbury visited the University of Amsterdam to perform work related to this showcase.

7.9 Contribution by INRIA-Parole

Researchers involved

Jacques Feldmar and Jonathan Ponroy

Activities & Achievements

- Activities: A first version of the INRIA phoneme recognition engine and of the KTH talking head has been integrated. A communication protocol has been implemented and tested. It is TCP-IP based and guaranties synchronization between the two processes. INRIA has integrated Gaussian Mixture Models (GMM) into its Voice Activation Detection (VAD). Quality of speech-non speech has been improved but still needs to be increased. The current VAD engine is based on an automaton. The topology and parameters of this automaton have been changed and improved. KTH has integrated some articulators to its talking head such as the tongue. Different display schemes are being investigated. A first version of the graphic user interface has been developed. It shows articulators as a vertical slice. - Achievements: A first version of the articulatory talking head driven by speech recognition has been developed.

Events

Jacques Feldmar Attended the Muscle meeting in Budapest on April 23 to present the first version of the articulatory talking.

Publications

"Evaluation of a talking head for helping HOH people in the classroom", Lorène Mourot, Marie Rovert, thesis for the degree of Speech therapist.

7.10 Contribution by IRIT-UPS

Researchers involved

Eduardo Sanchez-Soto, Julien Pinquier, Philippe Joly, Thomas Foures, Frédéric Gianni.

Activities & Achievements

The scientific activities provided for the showcases are described inside the Workpackages WP2 and WP5. Here are given the degree of achievements: Showcase ACADI: the interface evolves, with in particular: - navigation in the segmentation and in the video, - concatenation of segments to play and to see all the interventions of a person, - incorporation of DIVA (client/server video manager) in the system, - translation of the diarization source code (matlab ->C) - adaptation of each tool (face, speaker and costume detections) in an independent way. Showcase CASEWP with the development odf the web portal : - implementation of the interface - verification service of the submission validity - evaluation service of submissions - integration inside the plateform OSIRIM

Events

Participation to the Budapest Meeting (April 23th): State of the art of the ACADI showcase: presentation of the fusion method and demonstration of the current system (principal functionalities). Demonstration with the first version of the CASEWP portal

7.11 Contribution by IBaI

Researchers involved

Petra Perner, Horst Perner

Activities & Achievements

A presentation for the showcase at CeBIT has been done as well as a demo for the demonstration of the catalogue-based classifier. The Muscle sponsored conference on Mass Data Analysis of Images and Signals in Medicine, Biotechnology and Chemistry has been continuously prepared. The reviewing process has been handled. The advertisement among scientists and industry has been made. The conference proceeding is prepared.

Events

CeBit in Hannover
MDA 2007 www.mda-signals.de

7.12 Contribution by BILKENT UNIVERSITY

Researchers involved

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

Activities & Achievements

We have contributed with our Multi-modal Method for Detecting Fight among People at Unattended Places to the Flexible Test-Bed for Unusual Behavior Detection paper which will appear in ACM International Conference on Image and Video Retrieval 2007. There are several different algorithms produced in for visual surveillance and activity analysis. To obtain more robust systems it is desirable to integrate the different algorithms and fuse their results. In order to achieve this goal, István et al. from SZTAKI developed a flexible, distributed software collaboration framework and presented a prototype system for automatic event analysis. We are working with SZTAKI to integrate our Multi-modal Method for Detecting Fight among People module into their test bed. Recently, intelligent video analysis systems capable of detecting humans, cars etc were developed. Such systems mostly use HMMs or SVMs to reach decisions. They detect important events but they also produce false alarms. It is possible to take advantage of other low cost sensors including audio to reduce the number of false alarms. Most video recording systems have the capability of recording audio as well. Analysis of audio for intelligent information extraction is a relatively new area.

Automatic detection of broken glass sounds, car crash sounds, screams, increasing sound level at the background are indicators of important events. By combining the information coming from the audio channel with the information from the video channels, reliable surveillance systems can be built. In our module, a surveillance subsystem that utilizes both video and audio to detect fight among people at unattended places were developed. First, moving objects in video are segmented from the scene background by using an adaptive background subtraction algorithm and then segmented objects are classified into groups like human and human group using a silhouette based classification method. By analyzing the motion of the human groups and at the same time detecting screams or increasing sound in audio a decision is given to detect fight.

Publications

Petras et al. "Flexible Test-Bed for Unusual Behavior Detection", to appear in proceedings of CIVR'07, July 9-11, 2007, Amsterdam.

Researchers involved

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

Activities & Achievements

We have contributed with our Fire and Smoke Detection Method to the Dynamic Texture Detection, Segmentation and Analysis paper which will appear in ACM International Conference on Image and Video Retrieval 2007. 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 work together with S.Fazekas and D.Chetverikov from SZTAKI, Hungary and T.Amiaz and N.Kiryati from Tel Aviv University, Israel, we develop real-time dynamic texture detection methods in video and present solutions to video object classification based on motion information. Regarding our method, conventional point smoke and fire detectors typically detect the presence of certain particles generated by smoke and fire by ionisation or photometry. An important weakness of point detectors is that they are distance limited and fail in open or large spaces. The strength of using video in fire detection is the ability to monitor large and open spaces. Current fire and flame detection algorithms are based on the use of color and motion information in video. In this work, we not only detect fire and flame colored moving regions but also analyze the motion. Our experiments indicate that flame flicker frequency is not constant and it varies in time. In fact, variations in flame pixels can be considered as random events. It is wellknown that turbulent flames flicker with a frequency of around 10 Hz. Therefore, fire detection scheme was made more robust by detecting periodic high-frequency behavior in flame colored moving pixels compared to existing fire detection systems. Therefore a Markov model based modeling of flame flicker process produces more robust performance compared to frequency domain based methods. This is also valid for video based smoke detection.

Publications

Toreyin et al. "Dynamic Texture Detection, Segmentation and Analysis", to appear in proceedings of CIVR'07, July 9-11, 2007, Amsterdam.

Researchers involved

All showcase leaders and researchers who participated to the meeting.

Activities & Achievements

- The showcase project demonstration meeting was organized in Budapest on 23, 24 April. All showcase leaders and researchers participated to the meeting and presented the progress of their showcases. The Steering Committee was satisfied with the progress made.

8 MUSCLE Effort Table

Person-Months report for period : March - April 2007

Institute	WP1a	WP1b	WP2	WP3	WP4	WP5	WP6	WP7	Total
01 - ERCIM	2.2	0	0	0	0	0	0	0	2.2
03 - UCL	0	0	0	2	0	0	0	0	2
04 - KTH	0	0	0	1	2	1	0	0	4
05 - BILKENT	0	0.2	0.3	0	0.25	0.4	0.2	4.5	5.85
06 - VIENNA PRIP	0	0	0.3	3.7	0	0	0	2	6
07 - MTA SZTAKI	0	0	0	0.4	0	0	0.7	1.2	2.3
11 - TUG	0	0.09	0	2.12	0	0	1.85	0.18	4.24
12 - UPC	0	0	0.5	0.7	0	0	0	0.5	1.7
14 - UTIA	0	0.7	0	0.6	0.5	0.1	0.5	0.3	2.7
16 - AUTH	0	0	0	0	0.48	0.48	0	0	0.96
18 - TU VIENNA IFS	0	0	0	0	2	1.5	1	1	5.5
19 - ACV	0	0	0	0	0	0	0	0.14	0.14
21 - TECHNION-MM	0	0	0	0	0	0.11	0	0.28	0.39
22 - IBAI	0	0.01	0	0	0.34	0	0.76	0.33	1.44
26 - TAU-SPEECH	0	0	0	0.3	4	0	0	0	4.3
27 - TAU-VISUAL	0	0	0	2	0	0	0	1	3
32 - INRIA Ariana	0	0.23	0	0.62	0	0	0.26	0.38	1.49
32 - INRIA Parole	0	0	0	0	0	0	0	2	2
32 - INRIA Tex Mex	0	0	0	0	0	2.2	0	0	2.2
33 - GET	0	0	0	1.34	0	0	0	4.33	5.67
35 - UNIS	0	0	0	4.1	0	0.35	2	0	6.45
37 - ENSEA	0	0	0.1	0.25	0	0	0.35	0	0.7
38 - CNRS	0	0	0	0	1	0	0	0	1
39 - UPS – IRIT	0	0.2	1.3	0.5	1	1	1	2	7
42 - NUID / UCD	0	0.1	0	0	0	0	2.1	0	2.2
Total	**2.2**	**1.53**	**2.5**	**19.63**	**11.57**	**7.14**	**10.72**	**20.14**	**75.43**