



# MUSCLE

Network of Excellence

**Multimedia Understanding through Semantics, Computation and Learning**

Project no. FP6-507752

## **Bi-Monthly Progress Report #14: May – June 2006**

Due date of deliverable: 15.08.2006

Actual submission date: 20.08.2006

Start date of Project: 1 March 2004

Duration: 48 Months

**Name of responsible editor(s):**

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

Revision: 1.0

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

**Keyword List:**

## **Contents**

<b>1 Overview activities in WP 1</b>	<b>3</b>
<b>2 Overview activities in WP 2</b>	<b>4</b>
<b>3 Overview activities in WP 3</b>	<b>4</b>
<b>4 Overview activities in WP 4</b>	<b>5</b>
<b>5 Overview activities in WP 5</b>	<b>5</b>
<b>6 Overview activities in WP 6</b>	<b>11</b>
<b>7 Overview activities in WP 7</b>	<b>12</b>
<b>8 Overview activities in WP 8</b>	<b>13</b>
<b>9 Overview activities in WP 9</b>	<b>14</b>
<b>10 Overview activities in WP 10</b>	<b>15</b>
<b>11 Overview activities in WP 11</b>	<b>16</b>
<b>12 New MUSCLE Publications: Papers and Preprints</b>	<b>18</b>
<b>13 Resource Tables</b>	<b>19</b>

## 1 Overview activities in WP 1

### 1.1 General Scientific and Administrative Coordination

- Administrative and financial coordination of the network
- Preparation of JPA3-documents
- Organisation of the joint DELOS-MUSCLE summerschool on *Multimedia Digital Libraries: Machine Learning and Cross-modal Technologies for Access and Retrieval*, San Vincenzo, 12-17 June 2006 (in collaboration with Nozha Boujema, Valerie Gouet-Brunet). For more information, see below.
- Organisation of regular audio conferences
- Reimbursement of MUSCLE integration expenses (mobility support grants)
- Follow-up and posting of the deliverables.

### 1.2 DELOS-MUSCLE Summerschool

A lot of coordination effort went into the preparation, organisation and supervision of the first DELOS-MUSCLE summer school which took place in San Vincenzo (Italy) from 12-17 June 2006. For more information on the programme we refer to the summerschool website

<http://www-rocq.inria.fr/imedia/DelosMuscleSummerSchool2006/index.php>

The video-taped lectures are now available at the MUSCLE website:

[www.muscle-noe.org](http://www.muscle-noe.org) > dissemination > courseware.

There were 59 participants from 11 countries (for practical reasons related to classroom size, we had originally planned to limit the number to 50, but we relaxed that constraint slightly. A total of 22 students participated in the examination and obtained a certificate.

#### Geographical distribution:

- *European*: France 6, Germany 1, Greece 1, Hungary 2, Italy 39, Netherlands 1, Spain 3;
- *Non-European* : Algeria 1, Canada 1, India 2, Lebanon 1, Tunisia 1.

#### Network participation

- Number of students coming from DELOS/MUSCLE organizations: 21
- Number of students coming from other organizations: 38

**Summary of numerical scores on exit questionnaire** Non-percentage scores vary from 1 (poor), 2 (fair), 3 (good), 4 (very good) and 5 (excellent).

- 1 How would you rate the structure/organization of the event? **4.22**
- 2 Overall how effective were the speakers? **3.80**
- 3 Overall how effective did you find the printed slides? **4.04**
- 4 How well did you feel that the event addressed the area of multimedia digital libraries? **3.98**
- 5 Were professors available to interact with students and was their presence effective? **93%**  
Y/N

- 11 How do you rate this school in comparison to others or with similar events that you attended in the past? **4.00**
- 12 Did the fact that there was no school fee influence your decision to attend? **58%** y/n
- 18 How do you rate the quality of the San Vincenzo tower as a venue for courses of this kind? **3.93**
- 19 How did you rate the food? **2.93**
- 20 Please rate the social event. **4.31**
- 21 Please rate the Gala dinner. **4.33**

The responses to the other questions were in the form of free text and are difficult to summarize concisely.

## 2 Overview activities in WP 2

### 2.1 Contribution by CWI

**Researchers involved:** Eric Pauwels, Paul de Zeeuw, Margriet Brouwer

**Activities** General maintenance and updates of PHP/MySQL tools for information exchange and administrative and scientific reporting in WPs:

- Showcases and courseware;
- Mobility grants and reporting;
- Webform for report on internet-based collaboration;
- Preparation of hand-over of webtools to ERCIM who will supervise and maintain the website as of Sept 1, 2006.

## 3 Overview activities in WP 3

### 3.1 Contribution by TUVienna-PRIP

**Researchers involved:** Allan Hanbury,

**Activities** Organising MUSCLE evaluation activities in 2006. These include:

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

Two MUSCLE-sponsored workshops related to these activities are also being organised. The CIS competition result workshop will take place on the 11th of September in Berlin, Germany (before the DAGM conference). The MUSCLE/ImageCLEF 2006 workshop will take place on the 19th of September in Alicante, Spain (before the CLEF workshop).

Three programs were submitted to the coin competition, of which two worked. The tests are complete and two papers describing the methods have been submitted. An overview paper is under preparation. The MUSCLE/ImageCLEF workshop preparations are well underway. We have one invited speaker, Mark Everingham, co-organiser of the PASCAL Visual Object Challenge. Six papers have been submitted.

**Publications MP-codes:**

### 3.2 Contribution by CNR-ISTI

**Researchers involved:** Emanuele Salerno, Anna Tonazzini

**Activities** Document image database from Gerolamo Cardanos Contradicentium Medicorum (1663) partly online in the Software & Data Repository for E-team on unsupervised segmentation.

**Publications MP-codes:**

## 4 Overview activities in WP 4

### 4.1 Contribution by CNR-ISTI

**Researchers involved:** Ovidio Salvetti, Massimo Martinelli, Gabriele Pieri, Sara Colantonio, Davide Moroni, Patrizia Asirelli

**Activities** Participation to seminars and conferences: SWAMM 2006, WWW 2006, QIRT 2006, Muscle-Delos Summer School.

## 5 Overview activities in WP 5

### 5.1 Contribution by TUVienna-PRIP

**Researchers involved:** Allan Hanbury,

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

- Segmentation of images using the waterfall algorithm and watershed using volume extinction values on colour-texture gradients
- Matching of images using 2D colour histograms

Extensive experimentation on both of these topics has been done in the past 2 months, and 2 joint journal papers are under preparation.

**Publications MP-codes:** 181

### 5.2 Contribution by TUVienna-PRIP

**Researchers involved:** Allan Hanbury, Lech Szumilas, Branislav Micusik

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

- Collaboration with ARMINES-CMM on image segmentation and matching.
- Collaboration with KTH on texture features and scale detection.
- Collaboration with CEA on topics related to recognition of animals.

The following MUSCLE mobility grants have been used within this E-team (see the mobility grant reports for more specific descriptions of what was done during each visit):

- Lech Szumilas visited the CMM in October 2005 for a course on morphological segmentation.
- Allan Hanbury visited the CMM in April 2006 to continue work on segmentation and matching.
- Alireza Tavakoli Targhi (KTH) visited PRIP in March 2006. Work was done on texture features and automatic scale detection for animal recognition.
- Christophe Millet (CEA) visited PRIP in June 2006. The first experiments on animal recognition using Support Vector Machines were done.
- Lech Szumilas visited KTH in July 2006. Further work on texture features and animal recognition was done.

As a result of these exchanges, a number of joint papers are currently in preparation.

**Publications MP-codes:** 181

### 5.3 Contribution by TUVienna-PRIP

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

**Activities** Development of image segmentation algorithms which segment an image based on a sample of the texture to be found. This sample must be specified in advance by the user. The algorithm then attempts to mark all the regions in the image which correspond to the specified texture. It makes use of the minimum cut/maximum flow algorithm on a graph representation of the image. This problem is an instance of the one-class classification problem, as we have information on the texture to be located, but no information on the "background" (the rest of the image). This algorithm has been further developed to be fully automatic, yielding good results which will be presented at the ECCV 2006. Improving the efficiency of this method by creating a graph based on the results of a Maximally Stable Extremal Region (MSER) detector instead of directly on the pixels has been done and the results will be presented at the British Machine Vision Conference in September.

A texture detector for automatically locating textures in an image has been developed. It tries to find alternating colour patterns within an image using a clustering technique. The detected texture patches are then passed to the segmentation algorithm described above to attempt to locate the whole texture. The results of this work will be presented at the International Symposium on Visual Computing in November 2006.

Image segmentation by agglomerative hierarchical clustering in the CIELAB colour space has been investigated. To make this approach computationally feasible, an initial over-segmentation of the image was applied to reduce the number of colours in the image. A journal paper on this approach has been submitted.

An algorithm for classifying images into categories such as city/nature and inside/outside has been developed. It has been entered into the ImagEval evaluation campaign.

We have also investigated the keywords which have been used to annotate images in currently available image datasets. This has resulted in the creation of a combined keyword list.

**Publications MP-codes:** 200,204,218,291,292,472

### 5.4 Contribution by UCL

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

**Activities** Task 2: Image and Video Processing

## Sub-task 4: Saliency detection and visual features configuration

Work at UCL has concentrated on detecting new forms of saliency that are relevant to the understanding of visual content. A framework for attention mechanisms has been developed that enables different structures to be identified in images by incorporating transforms that correspond to the features being sought.

Reflective symmetries are detected using an attention mechanism in which pixel configurations are transformed through reflections before translation and checking for a match. Peaks in the distributions of reflection axis angles at which matches are found indicate the locations and strengths of the symmetries present in the image. The mid points of lines joining corresponding fork pixels lie along the axis of symmetry of the shape. Forks include some pixels that mismatch each other; this ensures that emphasis is given to image regions that contain attentive material. Papers were presented at ICAPR 2005 and MMSP 2005. In a similar fashion measures of perspective are computed using a scaling transform before testing for a match. Peaks in the distributions of matches across the image indicate the locations of vanishing points in the image. A paper has been submitted to ICIP 2006.

Earlier work on attention-base similarity has led to the development of a new algorithm that extracts colour correction parameters from pairs of images and enables the perceived illumination of one image to be imposed on the other. This time colour transforms are used as a means of obtaining the relative illumination between two images. The colour shift that stimulates the highest frequency of matches represents a measure of the relative illumination of the two images. We apply the reverse colour shift to pixels in the second image to obtain a transformation that approximates the illumination present in the first image. The work was presented at SPIE in January 2006 and an internet service has been established to evaluate usability and acceptability (<http://colourcorrection.bat.bt.co.uk/ColourCorrection/>).

Attention based methods for motion estimation are being investigated that do not depend upon specific features which are thought to characterise foreground objects or background regions. The technique maps attention mechanisms in time and generates motion vectors for each frame in a moving video sequence. Results have been compared with motion vectors derived from MPEG video encoders.

Whilst experimenting with real time DSP implementations of visual attention algorithms applied to video camera outputs, it has been observed that the highest measures of saliency in images are obtained at the point of best focus. This work promises to offer new techniques for optimising the informativeness of images during focusing operations or through the variation of other parameters e.g. spectrum of illumination, brightness, etc. A paper has been submitted to ICIP 2006.

Visual attention algorithms have also been used to guide the selection of seed points for region growing in image segmentation. Points of low attention are normally present in large areas of self-similar background regions that are easy to segment, whereas high attention regions are often in the vicinity of rapidly changing features where segmentation needs to be constrained.

Several findings from research on Content Based Image Retrieval have been published that have used attention based similarity measures. Most significantly attributes derived from a combination of contextual metadata and image similarity have been used to indicate the location at which camera phone images were taken. In this work the problem of image classification by grouping images into visual clusters and combining this with contextual metadata was undertaken with some success using a very diverse set of images. This work was presented at SPIE in January 2006.

Personnel exchanges with INRIA are planned for the end of April during which ideas taken from saliency research will be applied to problems in copy detection.

**Publications MP-codes:** 206, 207, 220

## 5.5 Contribution by TUG

**Researchers involved:** Horst Bischof, Peter Roth, Helmut Grabner

**Activities**

1. We hosted the ECCV 2006.

2. The framework for observing static scenes to detect unknown objects using the online AdaBoost feature selection was updated and structured.
3. Experiments for developing a robust AdaBoost using statistical methods are running.

**Publications MP-codes:**

## 5.6 Contribution by ACV

**Researchers involved:** Herbert Ramoser, Csaba Beleznai, Julia Puckmayr

**Activities** The AdaBoost training algorithm has been changed in order to handle edge orientation histogram (EOH) features on intensity images as well as on corresponding frame difference images. Furthermore several AdaBoost trainings, using this so called motion features additionally, have been carried out. To evaluate the impact of parameter changes, an independent test data set has been generated and annotated. Results show that the choice of the negative training set is crucial, i.e. adding moving vehicle parts, moving vegetation and so on to the set of negatives to reduce false alarms.

**Publications MP-codes:**

## 5.7 Contribution by CNR-ISTI

**Researchers involved:** Anna Tonazzini, Luigi Bedini, Ercan Kuruoglu, Sergio Bottini, Stefania Gnesi, Francesco Bonchi

**Activities** Based on the statistical electrophoresis signal model previously formalized, an algorithm for basecalling in DNA sequencing has been developed and tested on real data.

**Publications MP-codes:**

## 5.8 Contribution by CNR-ISTI

**Researchers involved:** Anna Tonazzini, Francesco Bonchi, Stefania Gnesi, Ercan Kuruoglu, Sergio Bottini

**Activities** Published: A. Tonazzini, F. Bonchi, S. Gnesi, E. Kuruoglu, S. Bottini, Analysis and modelling of genomic data, *Ercim News*, vol. 64 (2006), pp. 59-60, 2006

**Publications MP-codes:** 370

## 5.9 Contribution by CNR-ISTI

**Researchers involved:** Anna Tonazzini,

**Activities** The activity on the statistical modelling of electrophoresis time series for basecalling in DNA sequencing has been presented at the Conference RECOMB 2006, Venice, 1-4 April, 2006. The related basecalling algorithm has been described in a paper submitted at the Conference ICDM 2006 - Workshop on Mass-Data Analysis of Images and Signals in Medicine, Biotechnology and Chemistry MDA2006.

**Publications MP-codes:**



## 5.10 Contribution by CNR-ISTI

**Researchers involved:** Ovidio Salvetti, Davide Moroni

**Activities** Development of a procedure for shape characterization and deformation analysis of 2D/3D deformable visual objects. We defined a reference dynamic model, encoding morphological and functional properties of an objects class, capable to analyze different scenarios in heart left ventricle analysis. The developed method is suitable for generalization to the analysis of periodically deforming anatomical structures, where it could provide useful support in medical diagnosis.

**Publications MP-codes:**

## 5.11 Contribution by UCL

**Researchers involved:** Fred Stentiford, Li Chen, Rob Shilston, Shijie Zhang

**Activities** Task 2: Image and Video Processing Sub-task 4: Saliency detection and visual features configuration. Work at UCL has continued to explore attention-based mechanisms.

- The detection of perspective in colour images has yielded good results and a paper is to be presented at ICIP 2006.
- Investigations of similar architectures for motion detection and estimation show that information on the geometry of the moving object may be extracted. Work is continuing with surveillance videos.
- An on-line internet demonstration of attention-based focusing has been set up using remote PTZ cameras across the world. Focusing commands are issued depending on the results of image analysis. A paper is to be presented at ICIP 2006. The work is currently directed at the optimization of other image parameters for image informativeness.
- Attention-based similarity has been applied to the problem of copy detection. The precision and recall results are encouraging. A paper on similarity is to be published in Pattern Recognition and the work on copy detection is to be reported at ICANN 2006 in the Special Session on Visual Attention Algorithms and Architectures for Perceptual Understanding and Video Coding. Further papers have been submitted to CVMP 2006 and IEEE Trans on Circuits and Systems for Video Technology.
- Li Chen visited INRIA in April to explore areas of mutual interest in video copy detection and make plans for exchanging data. A return visit to UCL is expected in September or October.

**Publications MP-codes:** 421, 422, 423, 424.

## 5.12 Contribution by CEA

**Researchers involved:** Moëllic Pierre-Alain, Christophe Millet, Adrian Popescu, Gregory Grefenstette, Patrick Hede

**Activities** Activity Works for the E-Team : Choosing Features for CBIR and Automated Image Annotation

- Test of features extraction using texture and colour information (Gabor and Local pattern)
- Evaluation of automatic recognition of Animals in Corel database using SVM

**Publications MP-codes:**

### 5.13 Contribution by MTA-SZTAKI

**Researchers involved:** Tamas Sziranyi, Csaba Benedek

**Activities** **Color models of shadow detection in video scenes**

We address the problem of appropriate modelling of shadows in color images. Regarding the transition between the background and shadow domains, we use an effective statistical distribution which is robust against several forthcoming illumination artifacts in the video scenes. The model has a low number of free parameters, which makes it to be appropriate for adaptive working. The main issue is to compare several well known color spaces with this model. We qualified the metrics both in color based clustering of the individual pixels and in the case of Bayesian foreground-background-shadow segmentation. Experimental results on real-life videos show that CIE  $L^*u^*v^*$  color space is the most efficient.

**Publications MP-codes:**

### 5.14 Contribution by TAU-visual

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

**Activities** Shape symmetry is an important cue for image understanding. In the absence of more detailed prior shape information, segmentation can be significantly facilitated by symmetry. However, when symmetry is distorted by perspectivity, the detection of symmetry becomes non-trivial, thus complicating symmetry-aided segmentation. We developed an original approach for segmentation of symmetrical objects accommodating perspective distortion. The key idea is the use of the replicative form induced by the symmetry for challenging segmentation tasks. This is accomplished by dynamic extraction of the object boundaries, based on the image gradients, gray levels or colors, concurrently with registration of the image symmetrical counterpart (e.g. reflection) to itself. The symmetrical counterpart of the evolving object contour supports the segmentation by resolving possible ambiguities due to noise, clutter, distortion, shadows, occlusions and assimilation with the background. The symmetry constraint is integrated in a comprehensive level-set functional for segmentation that determines the evolution of the delineating contour.

**Publications MP-codes:** 462

### 5.15 Contribution by TAU-visual

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

**Activities** We develop a novel variational approach for mutual segmentation of two images of the same object. The images are taken from different views, related by projective transformation. Each of the two images may not provide sufficient information for correct object-background delineation. The emerging segmentation of the object in each view provides a dynamic prior for the segmentation of the other image. The foundation of the proposed method is a unified level-set framework for region and edge based segmentation, associated with a shape similarity term. The dissimilarity between the two shape representations accounts for excess or deficient parts and is invariant to planar projective transformation. The suggested algorithm extracts the object in both images, correctly recovers its boundaries, and determines the homography between the two object views.

**Publications MP-codes:** 463

## 5.16 Contribution by TUVienna-IFS

**Researchers involved:** Thomas Lidy, Andreas Rauber

### Activities

- Developed Map of Mozart, using SOM-based technology and audio feature extraction
- Research on Audio Feature Extraction from online radio streams; applied unsupervised learning for the creation of Radio Station Maps and Profiles (Fingerprints) of Radio Stations
- Contributed a Paper ("Visually Profiling Radio Stations") and a Poster ("The Map of Mozart") to the Annual International Conference on Music Information Retrieval.

**Publications MP-codes:** 464

## 5.17 Contribution by UCL

**Researchers involved:** Fred Stentiford, Li Chen, Rob Shilston, Ade Bamidele

**Activities** Work has continued in collaboration with INRIA-IMEDIA on video copy detection. Following Li Chens visit to INRIA in April, a return visit by Julien Law-To is planned for early August during which an evaluation of several detection techniques will be carried out. Work on attention based similarity measures has also been applied to copy detection in still images. A paper on Near-Duplicate Image Matching has been submitted to CVMP 2006. A paper entitled Attention-Based Similarity has now been published online in the journal Pattern Recognition. The demonstration of attention based focusing is now working through local PTZ cameras as well as remote cameras that are connected to the internet. The photo colour correction technique reported earlier forms the basis of an award of an NCGE fellowship to Ade Bamidele. This will enable him to travel the US for 6 months and establish a case for a spin out company based on this technology.

**Publications MP-codes:** 465,467

## 5.18 Contribution by CNR-ISTI

**Researchers involved:** Anna Tonazzini, Luigi Bedini

**Activities** The activity on DNA basecalling and DNA microarray data analysis has been presented at the Scientific Meeting of the Ercim Working Group on Biomedical Informatics, within the 2006 Ercim General Meeting (Budapest, 31 May 2006).

**Publications MP-codes:**

# 6 Overview activities in WP 6

## 6.1 Contribution by AUTH

**Researchers involved:** Constantine Kotropoulos, Margarita Kotti, Emmanouil Benetos, Ignacio Martinez de Lizarrondo, Gustavo Martins, Jaime Cardoso

**Activities** Speaker turn detection

Task 3: Cross-modal Integration for Multimedia Analysis and Recognition

Subtask 3.1: Video Analysis and Integration of Asynchronous Time-evolving Modalities

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

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

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

**Publications MP-codes:** 235,283

## 7 Overview activities in WP 7

### 7.1 Contribution by CNR-ISTI

**Researchers involved:** Emanuele Salerno, Ercan Kuruoglu, Anna Tonazzini, Luigi Bedini

**Activities** Proposal for E-team on unsupervised segmentation presented at the WP7 Focus Meeting at Rocquencourt.

**Publications MP-codes:**

### 7.2 Contribution by CNR-ISTI

**Researchers involved:** Emanuele Salerno, Ercan Kuruoglu, Anna Tonazzini, Luigi Bedini

**Activities** E-team on unsupervised segmentation started at the Istanbul Plenary & Technical Meeting. Participants: ARMINES, Bilkent, CWI, KTH, MTA-SZTAKI, TCD, UTIA. Webspaces for collaborative work under construction.

Past activity on document analysis has been presented at the national symposium: IV Congresso nazionale di archeometria Scienza e beni culturali, Pisa, 1-3 Febbraio 2006.

**Publications MP-codes:** 369

### 7.3 Contribution by CNR-ISTI

**Researchers involved:** Emanuele Salerno, Anna Tonazzini, Luigi Bedini, Ercan Kuruoglu

**Activities** Software & Data Repository (<http://docshare.isti.cnr.it/si/login.php>) and Web Bulletin Board (<http://mx3.isti.cnr.it/phpbb2/index.php>) for E-team on unsupervised segmentation ready to be used for collaborative work.

Published: A. Tonazzini, E. Salerno, L. Bedini, Fast correction of bleed-through distortion in grayscale documents by a Blind Source Separation technique, IJDAR, published online 9 March 2006

**Publications MP-codes:**

## 7.4 Contribution by CNR-ISTI

**Researchers involved:** Emanuele Salerno,

**Activities** A study on unsupervised separation-segmentation has been conducted for the case where the source patterns are statistically dependent. The approach followed is the maximization of non-gaussianity of the output processes. The application at present under study is remote-sensed image analysis.

"Dependent component analysis as a tool for blind spectral unmixing of remote-sensed images", by C. Caiafa, E. Salerno, A. Proto, and L. Fiumi, accepted for presentation to Eusipco 2006 Conference (Florence, September 2006).

**Publications MP-codes:**

## 8 Overview activities in WP 8

### 8.1 Contribution by CNR-ISTI

**Researchers involved:** Ovidio Salvetti, Sara Colantonio

**Activities** Method for unsupervised/supervised analysis of complex colour images. In particular, design and development of an automatic procedure for segmenting lymphatic cell nuclei represented in microscopic specimen colour images. The proposed method follows a two-step approach to, firstly, find the nuclei and, then, to refine the segmentation by means of a neural model, able to localize the borders of each nucleus. Experimental results have shown the feasibility of the method.

**Publications MP-codes:**

### 8.2 Contribution by AUTH

**Researchers involved:** Constantine Kotropoulos, Dimitrios Ververidis

#### **Activities Feature selection based on mutual correlation**

Feature selection is a critical procedure in many pattern recognition applications. There are two distinct mechanisms for feature selection namely the wrapper methods and the filter methods. The filter methods are generally considered inferior to wrapper methods, however wrapper methods are computationally more demanding than filter methods. A novel filter feature selection method based on mutual correlation is proposed. We assess the classification performance of the proposed filter method by using the selected features to the Bayes classifier. Alternative filter feature selection methods that optimize either the Bhattacharyya distance or the divergence are also tested. Furthermore, wrapper feature selection techniques employing several search strategies such as the sequential forward search, the oscillating search, and the sequential floating forward search are also included in the comparative study. A trade off between the classification accuracy and the feature set dimensionality is demonstrated on both two benchmark datasets from UCI repository and two emotional speech data collections.

**Publications MP-codes:** 473

### 8.3 Contribution by UU

**Researchers involved:** Niall Rooney, David Patterson, Mykola Galushka

**Activities** We published the following papers on the SOPHIA system related to its scalability as a document clustering mechanism and to the use of relevance feedback. We shown that SOPHIA is an efficient document clustering algorithm and that the use of static clustering facilitates an effective relevance feedback mechanism requiring only few relevance judgements.

**Publications MP-codes:** 460, 461

### 8.4 Contribution by TUVienna-IFS

**Researchers involved:** Thomas Lidy, Andreas Rauber

**Activities** We wrote Book chapter for MUSCLE WP8 Book project: Machine Learning Techniques for Multimedia Content

**Publications MP-codes:**

### 8.5 Contribution by TUG

**Researchers involved:** Horst Bischof, Helmut Grabner, Peter Roth, Thomas Mauthner

**Activities**

1. We hosted ECCV 2006.
2. Training from unlabeled data with an automatic initialisation using a combination of PCA and a MSER-tracker was implemented in C++ and is now running in real-time.
3. The robust active shape models were ported to C++.
4. We are experimenting with tracking players of a volleyball team using our developed AdaBoost tracking and particle filters respectively.

**Publications MP-codes:**

## 9 Overview activities in WP 9

### 9.1 Contribution by CNR-ISTI

**Researchers involved:** Ovidio Salvetti, Massimo Martinelli, Gabriele Pieri, Patrizia Asirelli, Michele Lunardi, Enrico Luperini, Marco Tampucci

### Activities

- Implementation of part of the java-based Infrastructure for MultiMedia Metadata Management (4M): implementation of mpeg-7 extractors from audio, image and video started (audio completed); Implementation of an XML database for mpeg-7 metadata based on eXist XML.
- Participation at the SWAMM 2006 workshop (22nd May, WWW2006 Edimburgh): <http://image.ntua.gr/swamm2006/resources/paper01.pdf>
- WWW2006 Edimburgh: on 24th May we attended the ERCIM Stand at the conference, presenting the activities of the MUSCLE NoE.

### Publications MP-codes:

## 9.2 Contribution by IBAI

**Researchers involved:** Petra Perner, Horst Perner

**Activities** Preparation of the workshop CD for the MUSCLE sponsored event Mass Data Analysis of Images and Signals.

### Publications MP-codes:

## 10 Overview activities in WP 10

### 10.1 Contribution by TSI-TUC

**Researchers involved:** Alex Potamianos, Manolis Perakakis, Michalis Toutoudakis

**Activities** Multi-Modal Dialogue Systems: Modality Selection

The emergence of powerful mobile devices such as personal digital assistants (PDAs) and smart-phones raises new design challenges and constraints that could be better addressed by a combination of more than one modalities such as speech and visual modalities. Combining multiple modalities efficiently is a complex task and requires both good interface design and experimentation to determine the appropriate modality mix. Issues such as device size, interface latency, speech recognition accuracy can seriously affect the efficiency and naturalness of the multimodal interface and bias the optimum modality mix. Few guidelines exist for selecting the appropriate mix of modalities. It is established that the visual modality is more efficient than speech while speech is the more natural interaction mode. However, it is often the case when designing multimodal user interfaces, that the developer is biased either towards the voice or the visual modalities. This is especially true if the developer is voice-enabling an existing graphical user interface (GUI)-based application or building a GUI for an existing voice-only service.

Our goal is to follow an approach that respects both modalities, creating an interface that is both natural and efficient. Towards that goal we implement a modality-selection multimodal system that is a mixture of the click-to-talk and open-mike multimodal modes. Modality-selection implements a simple version of the adaptive modality tracking algorithm introduced in Potamianos et al 2003. The unimodal and three multimodal systems are built using the Bell Labs Communicator platform and are evaluated on a travel reservation task. Our focus in this work, is on creating and evaluating various multimodal interaction modes in order to gain valuable experience in the area of modality integration and selection.

Recently, we evaluated two unimodal and three multimodal travel reservation systems on the desktop and PDA environments. Our evaluation experiments outlined some basic facts of multimodal dialogue system design: (1) Synergies between the speech and visual interaction modes exist in multimodal interfaces; the systematic modeling of these synergies requires further research. (2) When changing the relative efficiency of the input modes in multimodal interfaces, user input mode usage also changes.

(3) It is not always true that a multimodal (speech and visual) interface is more efficient or preferable to the unimodal interface. The “best” interface is both a function of relative unimodal interfaces efficiency and user usage.

Future work will focus on evaluating the unimodal and multimodal systems for varying levels of task complexity and unimodal interface efficiency (e.g., different speech recognition error levels). Through these experiments multiple measurement points for mode usage, unimodal and multimodal interface efficiency will be obtained; these results will help us better understand the relationship between efficiency, user satisfaction and input mode usage. By incorporating this knowledge into the multimodal dialogue system design process we aim at building adaptive multimodal interfaces that are natural, efficient and outperform traditional unimodal interfaces.

Bibliography:

- M. Toutoudakis, M. Perakakis, and A. Potamianos, “Mode Selection in Multimodal Dialogue Systems”, invited paper to Internat. Conf. on Intelligent Systems And Computing: Theory And Applications, Cyprus, July 2006.
- M. Perakakis, M. Toutoudakis, and A. Potamianos, Modality selection for multimodal dialogue systems, in Internat. Conf. on Multimodal Interfaces, (Trento, Italy), Oct. 2005.
- A. Potamianos, E. Ammicht, and E. Fosler-Lussier, Modality tracking in the multimodal Bell Labs Communicator, in Proc. Automatic Speech Recogn. and Underst. Workshop, (St. Thomas, U.S. Virgin Islands), Dec. 2003.

**Publications MP-codes:**

## 10.2 Contribution by UCL

**Researchers involved:** Fred Stentiford, Wole Oyekoya

**Activities** This period has been devoted to consolidation and analysis. A significant conclusion from the most recent results confirms that pre-attentive vision is used in visual search. Participants are able to find target images in sessions where images were viewed for periods of 200ms and 300ms. In both cases participants reported that they felt that they had little control over the interface, but in spite of this the targets were located faster than a random search. The work is currently being written up. A paper entitled Eye Tracking: A New Interface for Visual Exploration is to be published in the BT Technology Journal.

**Publications MP-codes:** 468

## 11 Overview activities in WP 11

### 11.1 Contribution by Bilkent University

**Researchers involved:** Yigithan Dedeoglu, Ugur Gudukbay, Ibrahim Demir, A. Enis Cetin

**Activities** A 3D human pose reconstruction framework is developed that works on single view video. 3D human pose reconstruction is a popular research area since it can be used in various applications. Currently most of the methods work for constrained environments, where multi camera views are available and camera calibration is known, or a single camera view is available, which requires intensive user effort. However most of the existing video data do not satisfy these constraints, thus they cannot be processed by these algorithms. In this study a framework is proposed to reconstruct 3D pose of a human for animation from a sequence of single view video frames.



The framework for pose construction starts with background estimation. Once the image background is estimated, the body silhouette is extracted by using image subtraction for each frame. Then the body parts in the silhouette are automatically labeled by using a 2D model-based approach. Finally, the 3D pose is constructed from the labeled human silhouette by assuming orthographic projection and then extracted parameters are applied to the model by using forward kinematics to produce animation. The proposed approach does not require camera calibration. The proposed framework assumes that the input video has a static background and it has no significant perspective effects and the performer is in upright position.

**Publications MP-codes:**

## 11.2 Contribution by TUVienna-IFS

**Researchers involved:** Thomas Lidy, Andreas Rauber

**Activities** eTeam: Content Analysis Showcase

- Collected data (video recordings from TV from different countries, including different languages) from participating teams;
- Converted data to MPEG1 format;
- Extracted audio from video;
- Extracted features from audio streams;
- Provided data collection + features on MUSCLE CAS server (TUVienna-PRIP).

**Publications MP-codes:**

## 11.3 Contribution by TUVienna-IFS

**Researchers involved:** Thomas Lidy, Andreas Rauber

**Activities** started work for the eTeam: Semantic from Audio on the project: Dynamic optimization of DSP parameters in music rendering, using audio feature extraction: provided different feature sets + manual genre information + automatic genre classification

**Publications MP-codes:**

## 12 New MUSCLE Publications: Papers and Preprints

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

- **MP-424:** Stentiford, Fred; *Attention based similarity*
- **MP-460:** Rooney, Niall; Patterson, David; Galushka, Mykola; Dobrynin, Vladimir; *A scalable document clustering approach for large document corpora*
- **MP-461:** Rooney, Niall; Patterson, David; Galushka, Mykola; Dobrynin, Vladimir; *A relevance feedback mechanism for cluster-based retrieval*
- **MP-462:** Riklin-Raviv, Tammy; Sochen, Nir; Kiryati, Nahum; *Segmentation by Level sets and Symmetry*
- **MP-463:** Riklin-Raviv, Tammy; Sochen, Nir; Kiryati, Nahum; *Mutual Segmentation with Level Sets*
- **MP-464:** Lidy, Thomas; Rauber, Andreas; *Visually Profiling Radio Stations*
- **MP-465:** Zhang, Shijie; Stentiford, Fred; *An attention based method for motion detection and estimation*
- **MP-466:** Toutoudakis, Michail ; Perakakis, Manolis; Potamianos, Alexandros; *Mode Selection in Multimodal Dialogue Systems*
- **MP-467:** Chen, Li; Stentiford, Fred; *An attention based similarity measure for colour images*
- **MP-468:** Oyekoya, Wole; Stentiford, Fred; *Eye Tracking: A New Interface for Visual Exploration*
- **MP-470:** Tonazzini, Anna; Bonchi, Francesco; Gnesi, Stefania ; Kuruoglu, Ercan; Bottini, Sergio; *Analysis and modelling of genomic data*
- **MP-471:** Micusik, Branislav; Hanbury, Allan; *Template Patch Driven Segmentation*
- **MP-473:** Haindl, Michal; Somol, Petr; Ververidis, Dimitrios ; Kotropoulos, Constantine ; *Feature Selection Based on Mutual Correlation*

## 13 Resource Tables

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

## MUSCLE - Effort Table 14

Participant	WP01a	WP01b	WP02	WP03	WP04	WP05	WP06	WP07	TOTAL
01- ERCIM	2.03							0.30	2.33
02 - CWI			0.25	1.00	0.80		1.15		3.20
03 - UCL				0.35	0.25	0.80			1.40
04 - KTH			0.10	0.85	0.66				1.60
05 - BILKENT		0.20	0.40	0.35	0.25		0.40	1.60	3.20
06 - VIENNA PRIP			0.20	2.20	1.60				4.00
07 - MTA SZTAKI			0.60	0.19	0.12		1.00	0.30	2.20
08 - UU							1.10		1.10
09 - CNR-ISTI		0.10	0.35	0.05	0.05		0.15		0.70
10 - FT									
11 - TUG		0.16	0.35	0.30	0.25		1.25		2.31
12 - UPC			0.10	0.19	0.12				0.40
13 - UFR					0.10		0.25		0.35
14 - UTIA	0.10		1.00				1.30	0.20	2.60
15 - UVA						1.15			1.15
16 - AUTH				0.20	0.10			0.25	0.55
17 - CEA			0.25	0.35	0.20	0.20			1.00
18 - TU VIENNA IFS				0.95	0.65	0.60	0.45	0.30	2.95
19 - ACV				0.20	0.02				0.22
20 - TECHNION-ML							0.15		0.15
21 - TECHNION-MM				0.26	0.10	0.28		0.65	1.29
22 - IBAI		0.11	0.33				0.67	0.22	1.33
23 - ICCS		0.10	0.10	0.02	0.05	1.20		0.20	1.67
24 - TSI-TUC		0.10	0.04			1.65	0.03	0.40	2.22
25 - ARMINES		0.10		0.30	0.25				0.65
26 - TAU-SPEECH			1.25	0.20	0.14		0.55		2.14
27 - TAU-VISUAL				1.35	1.10				2.45
28 - SEIBERSDORF			0.65						0.65
29 - TCD		0.25	2.80	0.25	0.25		0.75		4.30
30 - FORTH				0.24	0.18	0.82	0.25	1.30	2.79
31 - VTT		0.18				0.44			0.62
32 - INRIA Ariana		0.35	0.20	0.01	0.05		0.35	0.35	1.31
32 - INRIA Imedia				0.03	0.05	1.00	0.30	1.00	2.38
32- INRIA Parole		0.10	0.70	0.08	0.03			0.20	1.10
32- INRIA Tex Mex						0.92			0.92
32 - INRIA Vista		0.05	0.15	0.50	0.40	0.02	0.30	0.45	1.87
33 - GET				2.50	2.01		1.30		5.81
36 - UCAM-DENG			0.55	0.15	0.10		0.60		1.40
34 - LTU			1.45						1.45
35 - UNIS			0.19	0.04	0.05	0.45	0.19	0.73	1.65
37 - ENSEA		0.22		0.10	0.10		0.30		0.72
38 - CNRS			0.25	0.50	0.50				1.25
39 - UPS – IRIT									
40 - EC3									
41 - UPMC									
42 - NUID / UCD									
Total	2.13	2.02	12.26	13.69	10.50	9.53	12.79	8.45	71.38