

ΠΕΡΙΟΧΕΣ ΘΕΜΑΤΩΝ ΓΙΑ ΔΙΠΛΩΜΑΤΙΚΕΣ

Οραση Υπολογιστών:

1. Μη-γραμμικά Φίλτρα για Αποκατάσταση (Restoration/Denoising) Εγχρωμων ή/και Κινούμενων Εικόνων.
2. Στατιστική Ανίχνευση Ακμών/Περιγραμμάτων με Επιπεδοσύνολα και Gestalt νόμους.
3. Στατιστική Ανάλυση/Σύνθεση Υφής με τυχαία Φράκταλς και προχωρημένα θέματα στοχαστικών ανελίξεων.
4. Ανάλυση Σχήματος σε πολλαπλές κλίμακες και Εφαρμογές σε Αναγνώριση Αντικειμένων με βάση το Σχήμα τους.
5. Κατάτμηση εικόνων με εξέλιξη καμπυλών και Ενεργά Περιγράμματα.
6. Κατάτμηση εικόνων με Γράφους και Ελαχιστοποίηση Ενεργειακών Συναρτησιακών.
7. Εκτίμηση οπτικής Κίνησης με εξέλιξη καμπυλών και ενεργά περιγράμματα και Εφαρμογή σε Αυτόματη Πλοήγηση κινούμενων Ρομπότ.
8. Θεωρία Στερέωσης με Γεωμετρία Πολλαπλών Οψεων.
9. Θεωρία Γεωμ. Διαφορικών Εξισώσεων και Επιπεδοσυνόλων (level sets) για Ο.Υ.
10. Θεωρία Markov/Gibbs Random Fields για ανάλυση εικόνων και όραση υπολογιστών.
11. Αναγνώριση Προσώπων με μη-γραμμικές μεθόδους Principal Component Analysis.
12. Εφαρμογές ενός από τα ανωτέρω θέματα σε 2Δ ή 3Δ Βιοϊατρικές εικόνες.
13. Εφαρμογές ενός από τα ανωτέρω θέματα σε προβλήματα Τηλεπισκόπησης.
14. Εφαρμογές των #9 ή #10 σε Ανακατασκευή Εικόνων από έργα Τέχνης.
15. Βελτίωση της τεχνολογίας JPEG με Wavelets/Πυραμίδες και μη-γραμμικά Scale-Spaces.
16. Ανάλυση Υφής και Χρώματος σε Βιοϊατρικές εικόνες ιστών για προβλήματα κατηγοριοποίησης ή κατάτμησης.

Επεξεργασία Φωνής-Γλώσσας και Πολυμεσικών Σημάτων:

1. Ανάλυση-Σύνθεση Φωνής με Μη-γραμμικά μοντέλα (Διαμόρφωσης, Χάος-Φράκταλς).
2. Μοντελοποίηση συστήματος Ακοής και Εξαγωγή Χαρακτηριστικών για αναγνώριση.
3. Multimodal Detection-Indexing Audio-Video Σημάτων.
4. Ανάλυση/Σύνθεση σημάτων Μουσικής με μοντέλα Διαμόρφωσης.
5. Ανάλυση Σκηνης Ακουστικών Γεγονότων στον 3Δ Χώρο.
6. Θέματα αυτόματης Αναγνώρισης Ελληνικής Γλώσσας.
7. Θέματα Πολυπλοκότητας σε υπολογιστικά μοντέλα γλώσσας.

Σήματα, Συστήματα, και Έλεγχος:

1. Ανάλυση/Πρόβλεψη Χρονοσειρών με Χαοτικά συστήματα και Φράκταλς με εφαρμογές σε οικονομικά συστήματα, ή τηλεπικοινωνιακά δίκτυα, ή βιολογικά συστήματα.
2. Μη-γραμμικά συστήματα Ελέγχου (max-plus άλγεβρα) με εφαρμογές σε οικονομικά συστήματα ή τηλεπικοινωνιακά δίκτυα.

Επιπρόσθετα Ειδικά Θέματα με Λεπτομερή Περιγραφή στην Ιστοσελίδα:

<http://cvsp.cs.ntua.gr/diplomatikes.shtm>

Καθηγητής Πέτρος Μαραγκός
E-mail: maragos@cs.ntua.gr

Επιπρόσθετα Ειδικά Θέματα με Λεπτομερή Περιγραφή

1. Title: **Hierarchical, Graph-based morphology -related Segmentation Methods.**

Abstract: Mathematical Morphology offers a rich framework to segment images and video sequences, amongst which the watershed transform paradigm has proved to be very powerful and effective. Its only drawback is the use of markers (seeds) which serve as segmentation starting points, since they determine the number of segmentation regions. Hierarchical segmentation appears to be a solution to overcoming the marker problem. The main idea of hierarchical segmentation is to produce a multiscale segmentation by progressive merging of regions either using extraneous knowledge or not. Using a morphological segmentation mechanism (watershed transform) an initial partition is produced. A similarity criterion is defined based on which “similar” regions are merged producing a single regions. The procedure is repeated in multiple scales until the desired scale is reached. The desired scale is defined based on the level of detail and segmentation demands of the specific application. There are many ways to represent the image partition hierarchy, the most common of which is the graph representation, such as Min/Max Tree, Binary Partition Tree and Minimum Spanning Tree.

An interesting computer vision thesis would be to implement one (or more) of the above region-based image representation algorithms and apply it to morphological image segmentation problems.

Information: P. Maragos and Natasa Sofou (sofou@cs.ntua.gr)

2. Title: **Optimal Fusion of Multiple visual Cues (contrast, size, texture, color) for Segmentation.**

Abstract: Image segmentation is a very hard computer vision task since it is application dependent. By the term application dependent we mean that we can not define a global segmentation algorithm since there are many different image categories that need to be processed in a special way depending on their characteristics, in order to lead to good segmentation results. Special image features, known as visual cues, that can be used in the segmentation process are contrast, colour, texture, object size, depth and motion. In recent years, many segmentation methods have been developed taking advantage of these characteristics, but there are hardly any segmentation methods that use a combination of the above characteristics in order to produce the desired image partition.

An interesting computer vision thesis would be to implement an algorithm that performs fusion of the most significant image cues: contrast, size, texture and colour.

Information: P. Maragos and Natasa Sofou (sofou@cs.ntua.gr)

3. Title: **Modern Methods For Robust Speech Processing**

Abstract: The physiology of human auditory system has inspired in many ways methods for speech analysis and recognition. Psychophysical experiments have indicated the incorporation of perceptual characteristics, leading to robust methods for speech processing and feature extraction. These ideas have been related with the slow varying modulation spectrum. On the other hand speech features of the modulation type based on the AM-FM dulation model has proved recently to have increased accuracy in speech recognition experiments. It would be challenging to incorporate characteristics used by perceptually inspired feature extraction algorithms (PLP, RASTA, J-RASTA) into Modulation type processing. Further experiment with novel feature sets for robust speech processing and evaluate these feature sets in simple speech recognition experiments.

Index Terms: PLP, RASTA, AM-FM Modulation model, HTK

Information: P. Maragos, Nassos Katsamanis (nkatsam@cs.ntua.gr), Vassilis Pitsikalis (vpitsik@cs.ntua.gr)

4. Title: **Nonlinear Methods for Robust Speech Processing and Recognition**

Abstract: There has been strong evidence for the existence of various nonlinear phenomena during speech production. Current state of the art speech analysis and feature extraction methods based on the linear source-filter model cannot account for such phenomena, like turbulence. Fractal measures, Lyapunov exponents and dynamical system inspired methods have been used to measure valuable information for phoneme classification and speech recognition even under adverse conditions. It is challenging to utilize fractal measures after processing the speech signals with appropriate filters, either in 1-d space or in multi-dimensional spaces. Furthermore, quantify the underlying signal complexity, and/or predictability by measuring variant invariant measures. Finally, possibly, apply the resulting feature extraction methods to simple recognition experiments over noisy databases (AURORA 2/3).

Index Terms: robust speech processing, nonlinear methods, fractals

Information: P. Maragos, Vassilis Pitsikalis (vpitsik@cs.ntua.gr), Nassos Katsamanis (nkatsam@cs.ntua.gr)

5. Title: **Event Detection in Audio Streams, specifically Speech-Music Segmentation and Classification**

Abstract: Tackling automatic content analysis of audio data is of major importance. Audio cues, either alone or integrated with information extracted from other modalities, may contribute significantly to the overall semantic interpretation of data. Event detection in audio streams is an important aspect within this framework. The concept of "event" corresponds to a noteworthy happening and is application dependent. For example, applause, chanting, laughter or alterations in the speech rate of the sport-caster may be regarded as events in a sports-video. Speaker changes, changes between various speech quality levels, between speech and silence, or speech and music are common events in the case of broadcast news.

Event detection in audio streams aims at delineating audio as a sequence of homogeneous parts each one identified as a member of a predefined audio class. Determination of such audio classes (e.g. speech, music, silence, laughter, noisy speech etc.) is the first step in the design of an event detection algorithm. Obviously, these classes are inferred from the specificity of the application. The selection and extraction of appropriate audio-features ensues and is probably the most significant phase. It is understood that the correct choice of these features favors the successful completion of the following phases. Proper exploitation of the extracted attributes leads to the segmentation of the audio stream and classification of the resulting parts.

Discriminating between speech and music segments in audio is a case that has attracted many researchers active in the area. Reviewing the state of the art and exploring novel approaches to cope with the problem is the main objective of the proposed diploma thesis.

Index Terms: audio classification, event detection, speech/music classification

Information: P. Maragos, Nassos Katsamanis (nkatsam@cs.ntua.gr), Vassilis Pitsikalis (vpitsik@cs.ntua.gr)

6. Title: Event Detection in Multimedia with Application to Sports' Video Summarization

Abstract: Scientific and technological progress nowadays facilitates the continuous accumulation of data. This fact yields unprecedented opportunities for the transition of our society to the prospect of a knowledge-driven society. Semantic manipulation of the deluge of information that is being accrued is a basic prerequisite towards this direction. Diversity and multimodality of this information pose additional challenges.

Event Detection in Multimedia in order to achieve indexing and summarization is an aspect of the aforementioned analysis. It is mainly motivated by the growing need of fast or selective visualization and indexing of TV broadcasts or videos. It could contribute significantly to efficient browsing of a video by picking out the main excerpts of the program or by removing all the irrelevant sequences according to the program theme.

Suitable exploitation of both audio and visual modalities contained in a sports' video could enable the extraction of a proper summary, e.g. containing the scoring opportunities of the opponent teams in a football match. This is the objective of the proposed diploma thesis.

Index Terms: content analysis, video summarization

Information: P. Maragos, Nassos Katsamanis (nkatsam@cs.ntua.gr), George Papandreou (gpapan@cs.ntua.gr)

7. Title: Automatic Large Vocabulary Continuous Speech Recognition System for the Greek Language

Abstract: Though research in the area of Automatic Speech Recognition (ASR) has reached significant milestones, the performance of current systems is far from being comparable to the speech recognition abilities of human beings. There is a number of key issues in the area that still remain open and make the task of building a complete ASR system that will work in real environments, quite challenging. Building such a system for the Greek Language, taking all its specialties into consideration, appears to be even more interesting.

The system basically will include separate components for speech processing, acoustic modeling and language modeling and of course already developed tools (HTK) may be used.

Index Terms: speech recognition, language modeling, acoustic models, HMMs, HTK

Information: P. Maragos, Nassos Katsamanis (nkatsam@cs.ntua.gr), Vassilis Pitsikalis (vpitsik@cs.ntua.gr)

8. Title: Statistical Context Analysis of Texts via Hidden Markov Models

Abstract: Context retrieval and summarization are emergent issues in the field of data mining. In parallel, statistical methods have been used extensively for natural language modeling in speech recognition systems. It would be challenging to view the issue of statistically modelling the characteristics (e.g. statistics of n-grams) of a certain text type in the context of statistical structure of language. This would consist a first low level approach towards the classification of different text types.

Index Terms: natural language processing, Hidden Markov Models, data mining, text classification

Information: P. Maragos, Nassos Katsamanis (nkatsam@cs.ntua.gr), Vassilis Pitsikalis (vpitsik@cs.ntua.gr)

9. Title: Statistical Face Detection in Images (Στατιστικός Εντοπισμός Προσώπων σε Εικόνες)

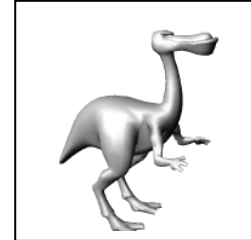
Abstract: Human face detection in still images is an important problem in Computer Vision, arising in important applications, like face recognition, area surveillance, visual lip-reading and Human Computer Interaction. Although the problem is well studied, it remains unsolved in its full generality, when many faces in various scales and orientations are allowed in the scene, the lighting conditions can vary widely and real-time performance is required. Among the face detection methods, the ones based on the statistical learning paradigm, requiring a training phase on an appropriate training set, are in the focus of current research and demonstrate the best performance. During this diploma project you will first review the relevant literature and then implement one or more of the state-of-art statistical face recognition algorithms. Methods for the quantitative evaluation of the performance of the implemented algorithms will also be explored.

Information: P. Maragos, George Papandreou (gpapan@cs.ntua.gr)

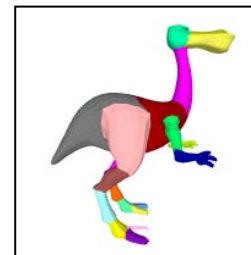
Διο (2) Θεματα Διπλωματικών σε Συνεργασία ΕΜΠ – «ΔΗΜΟΚΡΙΤΟΣ»

ΘΕΜΑ 1: Ιεραρχική Κατάτμηση 3Δ Γραφικών Μοντέλων

Η κατάτμηση 3Δ γραφικών μοντέλων αφορά στην αποσύνθεση πολύπλοκων αντικειμένων σε απλούστερα ή πρωτογενή 3Δ τμήματα στα οποία μπορεί να αποδοθεί σημασιολογική ερμηνεία (βλ. Σχήμα). Υπάρχει πληθώρα εφαρμογών στα γραφικά υπολογιστών που απαιτούν αναγκαίο βήμα την κατάτμηση 3Δ γραφικών μοντέλων. Χαρακτηριστικές εφαρμογές είναι: (i) η απλοποίηση γραφικού μοντέλου περαιτέρω συμπίεσή του; (ii) οι επικαλούμενες εφαρμογές “μεταμόρφωσης”, στις οποίες υπάρχει ανάγκη αντιστοίχισης επιμέρους δομών; (iii) η ανάκτηση 3Δ γραφικών μοντέλων, η οποία κάνει χρήση αναλλοίωτων περιγραφών που στη συγκεκριμένη περίπτωση αποδίδονται με ένα γράφο της τοπολογικής διευσθέτησης των επιμέρους τμημάτων; (iv) η ανίχνευση συγκρούσεων (collision detection), όπου η κατάτμηση βοηθά στον υπολογισμό των ιεραρχικά οριοθετημένων όγκων και (v) η απεικόνιση υφής (texture mapping).



ως
και



Αντικείμενο της προτεινόμενης διπλωματικής εργασίας αποτελεί ο σχεδιασμός, η ανάπτυξη καθώς και η αξιολόγηση αλγορίθμων που να αντιμετωπίζουν το πρόβλημα της ιεραρχικής κατάτμησης 3Δ γραφικών μοντέλων.

Οι προσεγγίσεις που θα υλοποιηθούν, θα βασιστούν στο μετασχηματισμό watershed που προέρχεται από τη Μαθηματική Μορφολογία και έχουν προηγούμενα εφαρμοσθεί σε 3Δ εικόνες.

Επιθυμητά χαρακτηριστικά υποψήφιου

- Να έχει παρακολουθήσει τα μαθήματα Ορασης υπολογιστή, Επεξεργασίας εικόνας και Γραφικά Υπολογιστών ;
- Γνώσεις προγραμματισμού σε C / C++, OpenGL.

Σχετική βιβλιογραφία

[Mangan-99] A. Mangan and R. Whitaker (1999), “Partitioning 3D surface meshes using watershed segmentation”, IEEE Transactions on Visualisation and Computer Graphics, vol. 5, no. 4, pp. 308-321.

[Page-03] D. Page, A.F. Koschan and M.A. Abidi (2003), “Perception-based 3D triangle mesh segmentation using fast marching watersheds”, In Proceedings of ICPR’03, pp. 27-32.

[Pratikakis-98] I. Pratikakis, “Watershed-driven image segmentation”, PhD Thesis, Vrije Universiteit Brussel.

[Zuckerberger-02] E. Zuckerberger, A. Tal and S. Shlafman (2002), “Polyhedral surface decomposition with applications”, Computers and Graphics, vol. 26, pp. 733-743.

Επικοινωνία: Καθ. Π. Μαραγκός, ΕΜΠ

και

Δρ. Ιωάννης Πρατικάκης

Tel:+30-210-6503183

Fax:+30-210-6532175

Εργαστήριο Υπολογιστικής Ευφυΐας

Ινστιτούτο Πληροφορικής και Τηλεπικοινωνιών

Εθνικό Κέντρο Φυσικών Ερευνών “Δημόκριτος”

15310 Αγία Παρασκευή, Αθήνα.

E-mail: ipratika@iit.demokritos.gr

ΘΕΜΑ 2: Κατάτμηση Αντικειμένων σε Εικονοσειρές (βίντεο)

Η κατάτμηση αντικειμένων σε εικονοσειρές (βίντεο) αποτελεί αντικείμενο μεγάλης πρόκλησης στις θεματικές περιοχές επεξεργασίας βίντεο και όρασης υπολογιστών. Αυτό επιπλέον δικαιολογείται από τις πολύ σημαντικές σύγχρονες εφαρμογές που απαιτούν ως θεμελιώδες συστατικό την κατάτμηση αντικειμένων σε εικονοσειρές. Χαρακτηριστικές εφαρμογές είναι: (i) η κωδικοποίηση εικονοσειρών βασισμένη σε αντικείμενα (π.χ. MPEG-4); (ii) ανάκτηση βίντεο βάσει περιεχομένου (π.χ. MPEG-7); (iii) παρακολούθηση από βίντεο για ασφάλεια και (iv) επεξεργασία βίντεο (video editing).

Αντικείμενο της προτεινόμενης διπλωματικής εργασίας αποτελεί ο σχεδιασμός, η ανάπτυξη καθώς και η αξιολόγηση αλγορίθμου που θα αντιμετωπίζει το πρόβλημα της κατάτμησης αντικειμένων σε εικονοσειρές.

Η προσέγγιση που θα υλοποιηθεί, θα λάβει υπόψη δύο επιμέρους συνιστώσες των οποίων ο συνδιασμός θα οδηγήσει στο τελικό αποτέλεσμα. Οι συνιστώσες αυτές είναι : (i) χωρική κατάτμηση εικόνας και (ii) εκτίμηση κίνησης βασισμένη στον περιορισμό της οπτικής ροής.

Οι επιμέρους συνιστώσες θα υποστηριχθούν από υπάρχουσες υλοποιήσεις [Pratikakis-03], [Pratikakis-99], [Pratikakis-98] στις οποίες θα βασιστεί ο υποψήφιος για να πετύχει το συνδιασμό τους.

Επιθυμητά χαρακτηριστικά υποψήφιου

- Να έχει παρακολουθήσει τα μαθήματα Ορασης υπολογιστή και Επεξεργασίας εικόνας ;
- Γνώσεις προγραμματισμού σε C / C++.

Σχετική βιβλιογραφία

[Erdem -04] C-E. Erdem, B. Sankur and A.M. Tekalp (2004), “Performance measures for video object segmentation and tracking”, IEEE Trans. on Image Processing, vol. 13, no. 7, pp. 937-951.

[Gatica-Perez-01] D. Gatica-Perez, C. Gu and M-T Sun (2001), “Semantic video object extraction using four-band watershed and partition lattice operators”, IEEE Trans. on CSVT, vol. 11, no. 5, pp. 603-618.

[Patras-04] I. Patras, M. Worring and R. Boomgaard (2004), “Dense motion estimation using regularization constraints on local parametric models”, IEEE Trans. on Image Processing, vol. 13, no. 11, pp. 1432-1443.

[Pratikakis-03] Pratikakis I., C. Barillot, P. Hellier and E. Memin (2003). Robust multiscale deformable registration of 3D Ultrasound images, International Journal of Image and Graphics – Special Issue on 3D Vision and Medical image Processing, vol. 3, no. 4 : 1-19.

[Pratikakis-99] Pratikakis I., Sahli H., and Cornelis J. (1999). Low level partitioning guided by the gradient watershed hierarchy, Signal Processing, 75(2):173-195

[Pratikakis-98] I. Pratikakis, “Watershed-driven image segmentation”, PhD Thesis, Vrije Universiteit Brussel.

Επικοινωνία: Καθ. Π. Μαραγκός, ΕΜΠ
και

Δρ. Ιωάννης Πρατικάκης Tel:+30-210-6503183

Fax:+30-210-6532175

Εργαστήριο Υπολογιστικής Ευφυΐας
Ινστιτούτο Πληροφορικής και Τηλεπικοινωνιών
Εθνικό Κέντρο Φυσικών Ερευνών “Δημόκριτος”
15310 Αγία Παρασκευή, Αθήνα
E-mail: ipratika@iit.demokritos.gr