



Ομάδα Όρασης Υπολογιστών, Επικοινωνίας Λόγου και Επεξεργασίας Σήματος
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
Εθνικό Μετσόβιο Πολυτεχνείο
<http://cvsp.cs.ntua.gr>

Ανάλυση Εικόνας και Όραση Υπολογιστών: Θεωρία και Εφαρμογές στην Αποκατάσταση Αρχαίων Τοιχογραφιών

Υποστήριξη διδακτορικής διατριβής

Γεώργιος Παπανδρέου

Περιεχόμενα διδακτορικής διατριβής

- Έναρξη διατριβής: Νοέμβριος 2003
- Ανάλυση εικόνας και όραση υπολογιστών
 - Ενδοσυμπλήρωση και κατάτμηση με τεχνικές μερικών διαφορικών εξισώσεων
 - Πολυκλιμακωτές αναπαραστάσεις σημάτων
 - Στατιστική μοντελοποίηση εικόνας στο πεδίο των κυματιδίων
 - Βασική εφαρμογή: Αποκατάσταση αρχαίων τοιχογραφιών
- Συνδυασμένη οπτική-ακουστική μοντελοποίηση φωνής
 - Οπτική-ακουστική μοντελοποίηση ομιλίας και τεχνικές για βέλτιστη σύμμειξη πολυτροπικών σημάτων
 - Μοντέλα όρασης υπολογιστών για περιγραφή προσώπου

(Μέρος I) Τοιχογραφίες στην Ανασκαφή Αιρωτηρίου Θήρας



Πρόγραμμα ΠΕΝΕΔ

- Συμμετέχοντα ιδρύματα και επικεφαλείς ερευνητές
 - ΣΗΜΜΥ - ΕΜΠ
 - Π. Μαραγκός (Ε.Υ.), Κ. Τζαφέστας
 - Ανασκαφή Θήρας
 - Χ. Ντούμας, Α. Βλαχόπουλος
 - Αρχιτεκτονική Σχολή ΑΠΘ
 - Κ. Παλυβού
 - Πανεπιστήμιο Ιωαννίνων, Τμήμα Ιστορίας & Αρχαιολογίας
 - Ι. Λώλος
- Υποψήφιοι διδάκτορες
 - Γ. Παπανδρέου (ΕΜΠ): Ψηφιακή αποκατάσταση τοιχογραφιών
 - Μ. Αληφραγκής (ΕΜΠ): 3-Δ καταγραφή & ανακατασκευή κτιρίων
 - Φ. Γεώργια (Ανασκαφή Θήρας & Πανεπιστήμιο Ιωαννίνων): Αρχαιολογική μελέτη τοιχογραφιών *κτιρίου Β*.
 - Κ. Αθανασίου (ΑΠΘ): Ψηφιακές τεχνικές στην αρχιτεκτονική μελέτη προϊστορικών μνημείων
- Χρηματοδότηση από τη ΓΓΕΤ (ΠΕΝΕΔ-03)
- Ιστοσελίδα: <http://cvsp.cs.ntua.gr/projects/AkrotiriThera>

Ανασκαφή Ακρωτηρίου Θήρας

- Προϊστορικός οικισμός Ακρωτηρίου Θήρας
 - Καταστροφή από έκρηξη ηφαιστείου το 17ο αιώνα π.Χ.
 - Ανασκαφές: Σ. Μαρινάτος (1967-74), Χ. Ντούμας (1974-σήμερα)
 - Ιδιαίτερα σημαντικά ευρήματα
- Τοιχογραφίες
 - Ποικίλος βαθμός διατήρησης/επίτιμο έργο συντήρησης
 - Σειρά προβλημάτων με έντονο επιστημονικό ενδιαφέρον
 - Επίκεντρο έρευνας: ψηφιακή αποκατάσταση τοιχογραφιών με ενδοσυμπλήρωση εικόνας



Αντιλόπες



Πυγμάχοι

Φωτογραφική αποτύπωση τοιχογραφιών

- Εργαστήριο τοιχογραφιών της ανασκαφής (Ξεστή 3 & Κτίριο Β)
 - Σεπτέμβριος 2006, Μάιος 2007, Σεπτέμβριος 2008



Ακολουθία
Ανδρών



Ακολουθία
Γυναικών



Κροκοσυλλέκτριες
& Πότνια



Καλαμιώνας

- Εθνικό Αρχαιολογικό Μουσείο (κτίριο Β)

- Νοέμβριος 2006



Αντιλόπες



Πυγμαχοί

Συνένωση φωτογραφικών στιγμιότυπων

- Σύνθεση φωτογραφικών μωσαϊκών υπερ-υψηλής ανάλυσης
- Υπολογιστική φωτογραφία



250 M-Pixels (150 dpi)



40 M-Pixels (70 dpi)



6 M-Pixels (30 dpi)

Πολυκλιμακωτή τεχνική φωτογράφισης



- ❑ 3 επίπεδα ανάλυσης
- ❑ 3 στάθμες έκθεσης



Διαδοχικά Στάδια Σύνθεσης Μωσαϊκών

- Λήψη φωτογραφιών
 - σχετική θέση κάμερας-σκηνής, πηγή φωτισμού
 - αλληλεπικάλυψη στιγμιοτύπων
- Ευθυγράμμιση εικόνων με εύρεση αντιστοιχιών
 - χαρακτηριστικά σημεία
 - επιπλέον περιορισμοί
- Συνένωση εικόνων και εξάλειψη ραφών
 - οπτικά «φαντάσματα»
 - διαφορές έκθεσης
- Διαδικασία διευκολύνεται από ελεύθερα διαθέσιμο λογισμικό ανοιχτού κώδικα

Αρχική καθολική άποψη



Ανορθωμένη καθολική άποψη



Αντιστοίχιση σημείων

The screenshot displays a software interface for point matching between two images of a damaged ancient wall painting. The left image is labeled '1. - DSC_0004.JPG' and the right image is '0. - DSC_0001.JPG'. Both images show a fragment of a wall with a painting of a figure and a red brick pattern. Numerous points are marked on both images with small colored squares and numbers (0-24). A table below the images lists the coordinates and alignment of these points. The table has columns for point number, left and right x and y coordinates, alignment, and distance. The distance for all points is 0.00. The interface also includes a 'Zoom' dropdown set to 'fit to window', a 'Fine-tune' button, and checkboxes for 'auto fine-tune', 'auto add', and 'auto estimate'. At the bottom, there are input fields for 'x' and 'y' coordinates and a 'mode' dropdown set to 'normal'.

#	left x	left y	right x	right y	Alignment	Distance
0	2020.66	1169.37	1008.11	586.51	normal	0.00
1	2742.11	363.74	1293.41	277.68	normal	0.00
2	1564.09	815.86	835.40	446.36	normal	0.00
3	1810.08	399.79	934.25	284.60	normal	0.00
4	1316.35	1040.50	736.99	535.56	normal	0.00
5	1316.35	1040.50	736.99	535.56	normal	0.00
6	222.22	1169.37	212.22	222.22	normal	0.00
7	222.22	1169.37	212.22	222.22	normal	0.00
8	222.22	1169.37	212.22	222.22	normal	0.00
9	222.22	1169.37	212.22	222.22	normal	0.00
10	222.22	1169.37	212.22	222.22	normal	0.00
11	222.22	1169.37	212.22	222.22	normal	0.00
12	222.22	1169.37	212.22	222.22	normal	0.00
13	222.22	1169.37	212.22	222.22	normal	0.00
14	222.22	1169.37	212.22	222.22	normal	0.00
15	222.22	1169.37	212.22	222.22	normal	0.00
16	222.22	1169.37	212.22	222.22	normal	0.00
17	222.22	1169.37	212.22	222.22	normal	0.00
18	222.22	1169.37	212.22	222.22	normal	0.00
19	222.22	1169.37	212.22	222.22	normal	0.00
20	222.22	1169.37	212.22	222.22	normal	0.00
21	222.22	1169.37	212.22	222.22	normal	0.00
22	222.22	1169.37	212.22	222.22	normal	0.00
23	222.22	1169.37	212.22	222.22	normal	0.00
24	222.22	1169.37	212.22	222.22	normal	0.00

SIFT descriptor: Lowe, Int. J. Comp. Vis, 2004

Ευθυγράμμιση λήψεων μεσαίας κλίμακας



Ευθυγράμμιση λήψεων λεπτομερούς κλίμακας



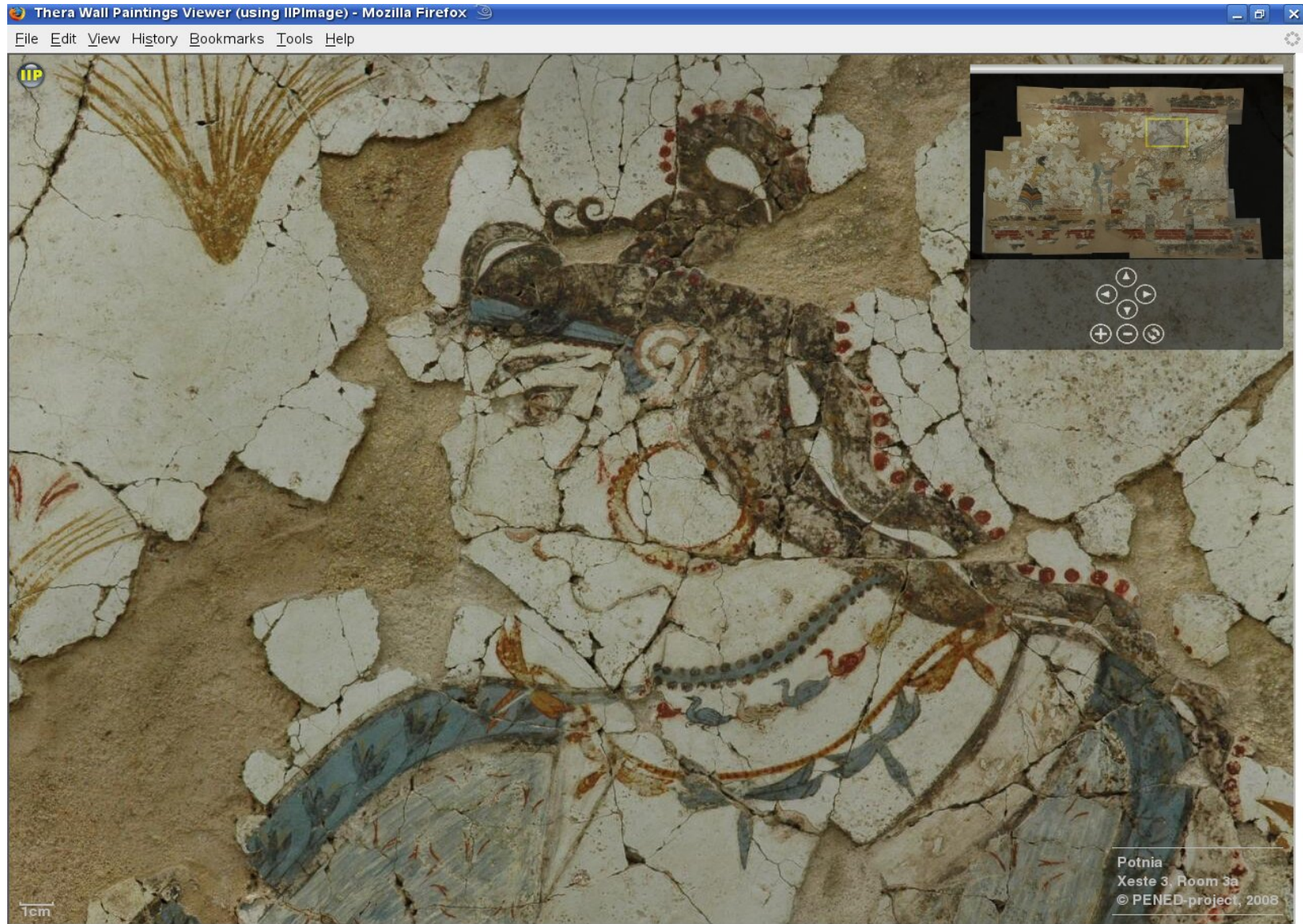
Ευθυγραμμισμένες λήψεις



Συρραφή μωσαϊκού

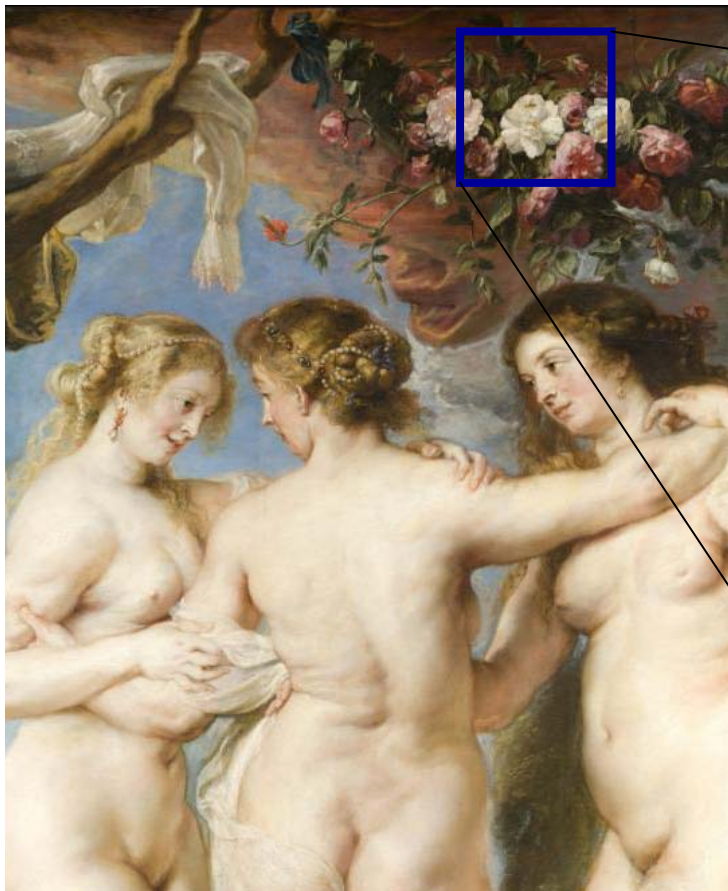


Πλοήγηση σε συνενωμένες φωτογραφικές συνθέσεις



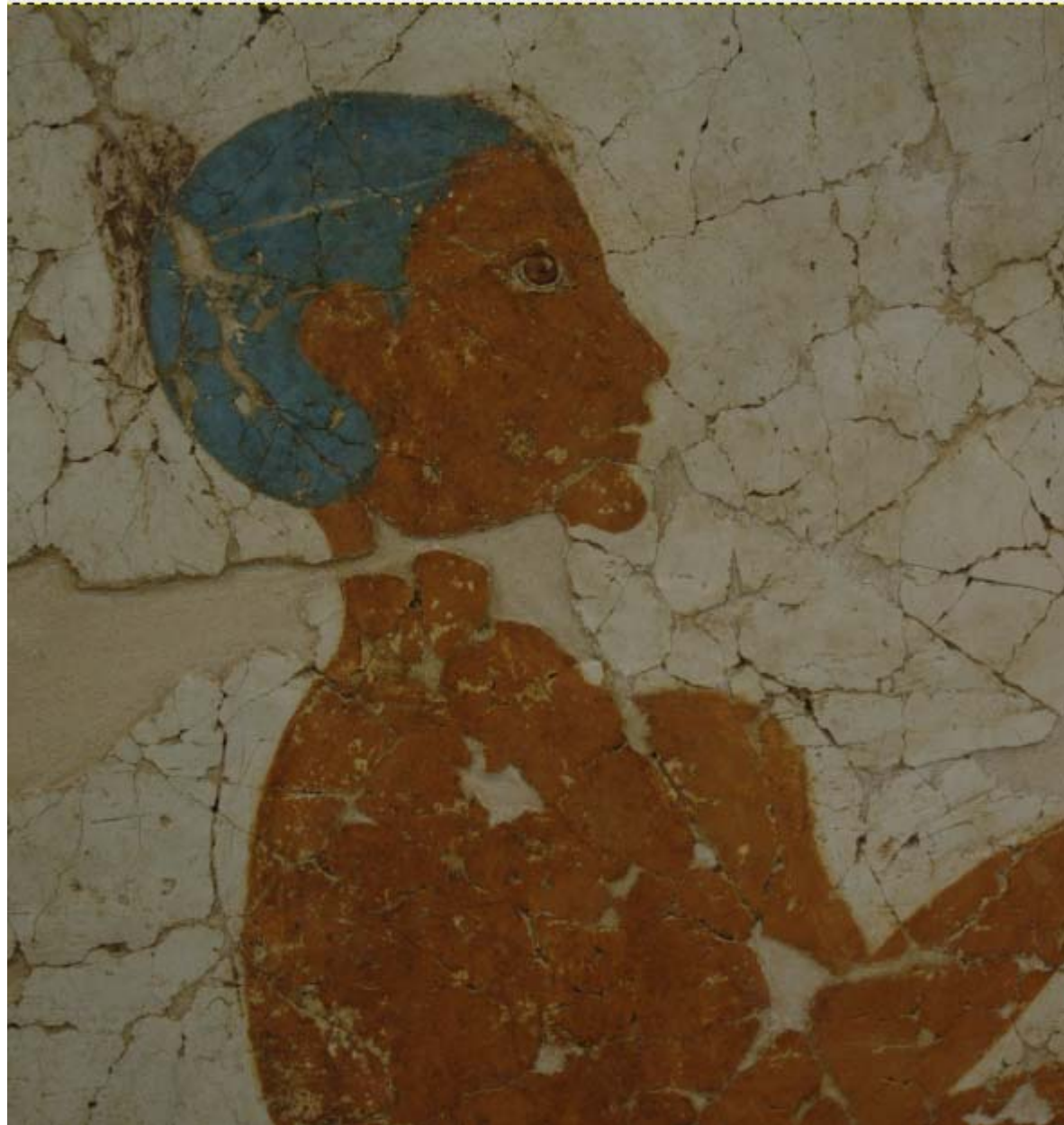
Πίνακες από το Prado στο Google Earth

- Φωτογραφίες (έκτασης gigapixel) διάσημων πινάκων του μουσείου Prado (Ιανουάριος 2009)



(Μέρος II) Ενδοσυμπλήρωση εικόνων



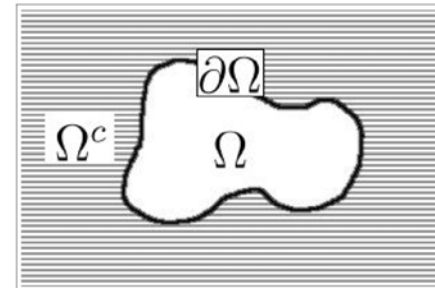






Ενδοσυμπλήρωση εικόνων

- Ενδοσυμπλήρωση (inpainting). Ορισμός:
Παρεμβολή τιμών σε ελλειπούσες περιοχές μιας εικόνας, δεδομένων των τιμών στις ορατές περιοχές της, με αξιοποίηση γνώσης σχετικά με τη συνήθη μορφή φυσικών εικόνων.

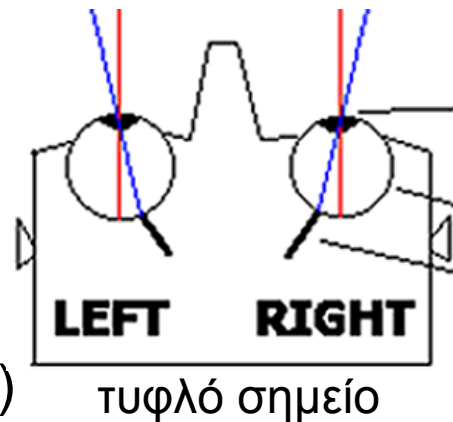


- Μοντέλα που ενσωματώνουν:
 - Διαδικασία αλλοίωσης της εικόνας (μέτρηση)
 - Κανονικότητες φυσικών εικόνων (πρότερη γνώση)
 - Συγκερασμό μέτρησης & πρότερης γνώσης (σύμμιξη)

Ενδοσυμπλήρωση & ανθρώπινη οπτική αντίληψη

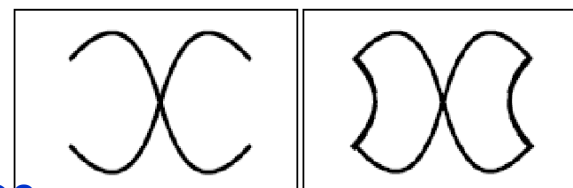
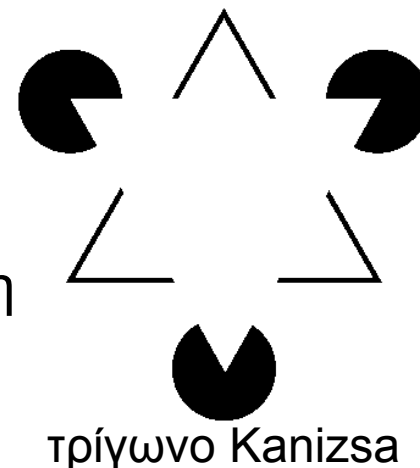
■ Εικόνες: 2-Δ όψεις του φυσικού κόσμου

- Οπτική αντιστροφή → έντονη αμφισημία
- Μηχανισμοί κανονικοποίησης
- Ασύνειδος συμπερασμός (Helmholtz, 19^{ος} αι.)



■ Ανθρώπινη όραση & συμπλήρωση πληροφορίας

- Τυφλό σημείο/σκοτώματα αμφιβληστροειδούς
- Απατηλά περιγράμματα/αντιληπτική οργάνωση
- Κατανόηση/οικονομική περιγραφή εικόνας
- Λειτουργίες χαμηλού-μέσου-υψηλού επιπέδου όρασης



κλείσιμο

Περιοχές εφαρμογής & σχετιζόμενες τεχνικές

■ Περιοχές εφαρμογής

- Κινηματογράφος: αποκατάσταση φιλμ σε ταινιοθήκες και ύστερη επεξεργασία ταινιών/ειδικά εφέ
- Επανάγγιμα φωτογραφιών
- Απόκρυψη σφαλμάτων μετάδοσης σε βιντεοροές
- Ψηφιακή διαφύλαξη πολιτισμικής κληρονομιάς

■ Συγγενή προβλήματα

- Σύνθεση υφής
- Ψηφιακή επαύξηση ανάλυσης εικόνας & αποψηφίδοποίηση
- Εκδοχές 1-Δ & 3-Δ (ενδοσυμπλήρωση σημάτων ήχου/βίντεο)

■ Σημαντικό θεωρητικό ενδιαφέρον

Ταξινόμηση τεχνικών ενδοσυμπλήρωσης

- Μη-παραμετρικές δεδομενοστρεφείς τεχνικές βασισμένες σε υποδείγματα
 - Εξανλητική αναζήτηση ομοειδών γειτονιών για «αντιγραφή & επικόλληση» εικονοστοιχείων ή μπλοκ της εικόνας

- Ιδιαίτερα ικανοποιητικά οπτικά αποτελέσματα

Popat and Picard (1997), De Bonet (1997), Efros and Leung (1999), Criminisi et al. (2003), ...

- Τεχνικές μεταβολικού λογισμού και ΜΔΕ

- Διάχυση περιεχομένου εικόνας κατά μήκος δομικών της στοιχείων

Masnou and Morel (1998), Bertalmio et al. (2000), Chan and Shen (2001), ...

Συνεισφορά: επιτάχυνση μεθόδων ΜΔΕ με πολυπλεγματικούς αλγορίθμους

- Παραμετρικά πιθανοτικά μοντέλα

- Περιγράφουν και αναπαράγουν βασικές στατιστικές ιδιότητες εικόνων

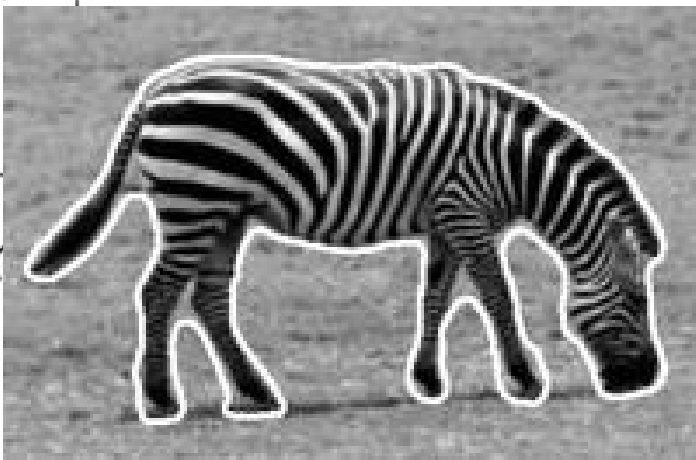
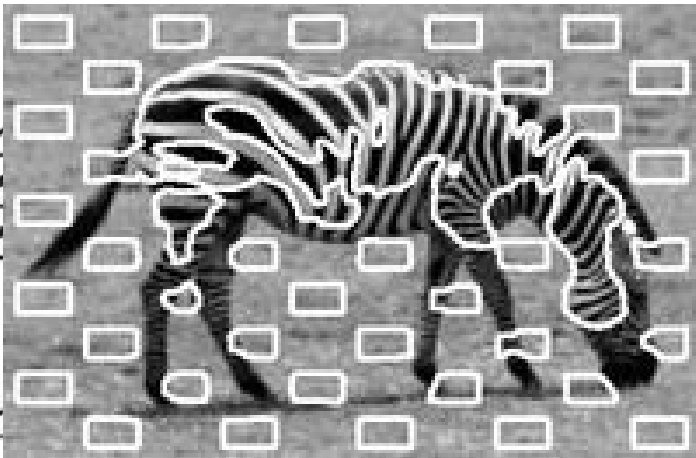
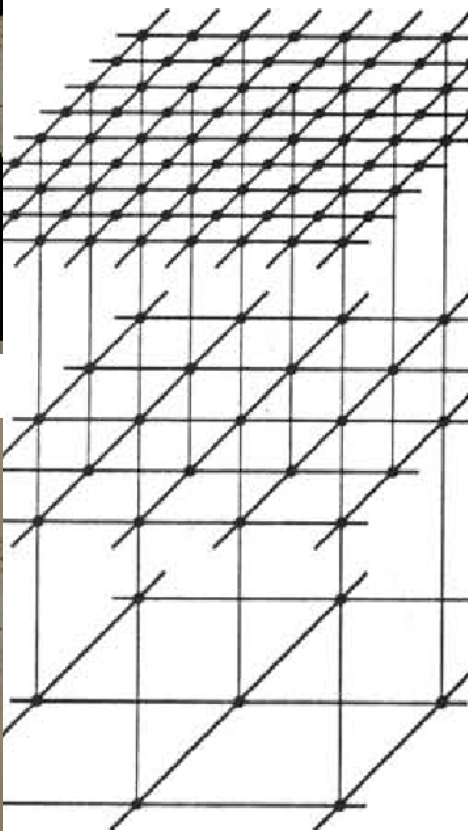
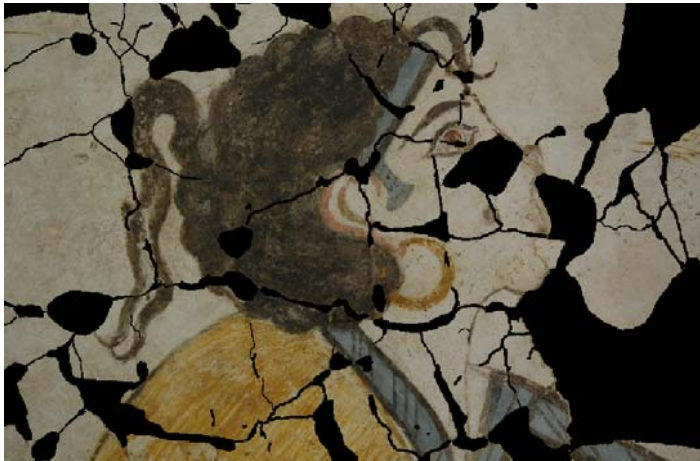
- Περιγραφή κυρίως στο πεδίο κάποιου μετασχηματισμού «κυματιδίων»

Zhu and Mumford (1997), Portilla and Simoncelli (2000), ...

Elad et al. (2005), Fadili and Starck (2005), Guleryuz (2006), Hua and Orchard (2007), ...

Συνεισφορά: βελτιωμένο πιθανοτικό μοντέλο στο πεδίο κυματιδίων που αξιοποιεί διακλιμακικές εξαρτήσεις μεταξύ συντελεστών κυματιδίων

(Μέρος III) Πολυπλεγματικοί Αλγόριθμοι για ΜΔΕ Ενδοσυμπλήρωσης & Κατάτμησης



Μοντέλα μερικών διαφορικών εξισώσεων στην όραση υπολογιστών: Χώροι κλίμακας

$$\frac{\partial u}{\partial t} = \text{div}(\nabla u)$$

$$\frac{\partial u}{\partial t} = \text{div}(g(\|\nabla u\|)\nabla u)$$

$$\frac{\partial u}{\partial t} = \text{div}\left(\frac{\nabla u}{\|\nabla u\|}\right)$$

Gaussian από γραμμική διάχυση



Μη-γραμμική ανομοιογενής διάχυση



Μη-γραμμική ολικής μεταβολής



ΜΔΕ και ελαχιστοποίηση ενέργειας για ενδοσυμπλήρωση

- Στενή σχέση ΜΔΕ & μεταβολικού λογισμού (Euler-Lagrange)

$$\min E(u) = \iint_{\Omega} \|\nabla u\|^2 + \mu(u - u_0)^2 dx \longrightarrow \frac{\partial u}{\partial t} = \operatorname{div}(\nabla u) - \mu(u - u_0)$$

όροι κανονικοποίησης όροι μέτρησης

$$\min TV(u) = \iint_{\Omega} \|\nabla u\| + \mu(u - u_0)^2 dx \longrightarrow \frac{\partial u}{\partial t} = \operatorname{div}\left(\frac{\nabla u}{\|\nabla u\|}\right) - \mu(u - u_0)$$

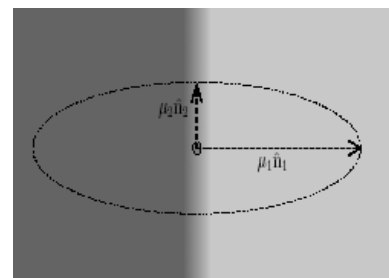
- Ενδοσυμπλήρωση-Κανονικοποίηση: Το είδος του συναρτησιακού καθορίζει τι θεωρούμε «τυπική εικόνα»
 - Ενέργεια «μεμβράνης»: ευνοεί ιδιαίτερα ομαλή ανακατασκευή
 - Συναρτησιακό ολικής μεταβολής: επιτρέπει ακμές στην εικόνα
- Ενδοσυμπλήρωση-Μέτρηση
 - $\mu = \mu(x, y)$ ανάλογα με ύπαρξη/πιστότητα παρατήρησης

ΜΔΕ ανισοτροπικής διάχυσης

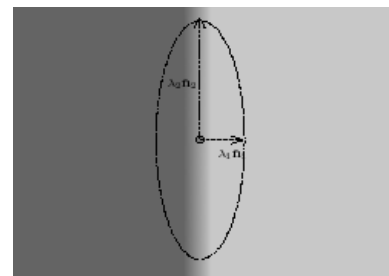
- Γενίκευση: από βαθμωτό συντελεστή διάχυσης σε τανυστή διάχυσης
 - Καλύτερη συνέχιση ακμών σε ενδοσυμπλήρωση (+)
 - Ψεύδεργα «ιμπρεσιονιστικού» τύπου (+/-)

$$\frac{\partial u}{\partial t} = \operatorname{div}(D\nabla u) - \mu(u - u_0)$$

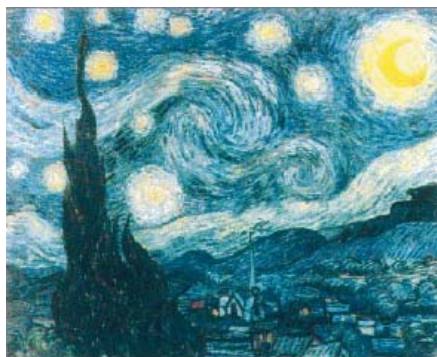
$$D(x, t) = \lambda_1 \hat{\mathbf{n}}_1 \hat{\mathbf{n}}_1^T + \lambda_2 \hat{\mathbf{n}}_2 \hat{\mathbf{n}}_2^T$$



δομικός
τανυστής



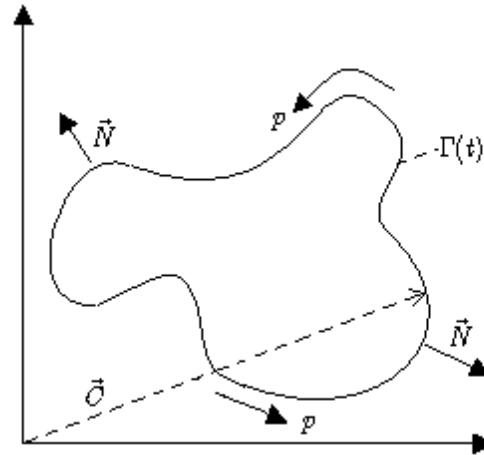
τανυστής
διάχυσης



Κατάτμηση με ενεργά περιγράμματα και αναπαράσταση επιπεδοσυνόλων

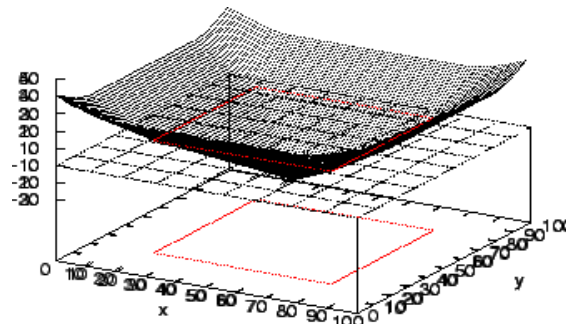
- Εξέλιξη καμπύλης

$$\frac{\partial \vec{C}(p,t)}{\partial t} = V \vec{N}(p,t)$$



- Εξέλιξη επιπεδοσυνόλου

$$\frac{\partial u(x,y,t)}{\partial t} = V \|\nabla u\|$$

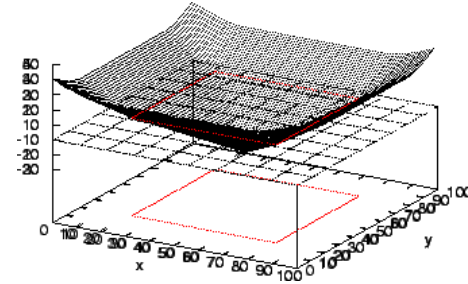


u συνάρτηση
εμβύθισης

Κατάτμηση με γεωμετρικά ενεργά περιγράμματα

- Γεωδαιτικά ενεργά περιγράμματα (ακμές):

$$\frac{\partial u}{\partial t} = |\nabla u| \operatorname{div} \left(g(|\nabla I|) \frac{\nabla u}{|\nabla u|} \right) + cg(|\nabla I|) |\nabla u|$$



u συνάρτηση
εμβύθισης

- Ενεργά περιγράμματα δίχως ακμές (περιοχές):

$$\frac{\partial u}{\partial t} = |\nabla u| \left\{ \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) + \frac{\alpha}{2} \left((c_R - I)^2 - (c_{R'} - I)^2 \right) \right\}$$

- Ανταγωνισμός περιοχών & γεωδαιτικές ενεργές περιοχές:

$$\frac{\partial u}{\partial t} = |\nabla u| \left\{ \operatorname{div} \left(g(|\nabla I|) \frac{\nabla u}{|\nabla u|} \right) + \alpha \log \frac{Pr(x|R)}{Pr(x|R')} \right\}$$

- Αρχικοποίηση σε μετασχηματισμό απόστασης

$$|\nabla u| = 1$$

Συγγενής μορφή ΜΔΕ ενδοσυμπλήρωσης/κατάτμησης

- Εξισώσεις γεωμετρικών ενεργών περιγραμμάτων

$$\frac{\partial u}{\partial t} = |\nabla u| \left\{ \underbrace{\operatorname{div} \left(g(|\nabla I|) \frac{\nabla u}{|\nabla u|} \right)}_{\text{εσωτερικές δυνάμεις}} + \underbrace{F(u)}_{\text{εξωτερικές δυνάμεις}} \right\}$$

- Εξισώσεις ανισοτροπικής διάχυσης

$$\frac{\partial u}{\partial t} = \underbrace{\operatorname{div}(D\nabla u)}_{\text{όρος διάχυσης}} - \underbrace{\mu(u - u_0)}_{\text{όρος αντίδρασης}}$$

- 1ος όρος (εσωτερικές δυνάμεις/διάχυση)

- Παραβολικού τύπου / Αριθμητικά άκαμπτος

- 2ος όρος (εξωτερικές δυνάμεις/αντίδραση)

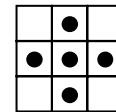
- Γραμμικός ή μη-γραμμικός / Αριθμητικά εύκαμπτος

Χωρική διακριτοποίηση

- Συνεχές πρόβλημα

$$\frac{\partial u(x, y, t)}{\partial t} = \text{div}(g(x, y)\nabla u)$$

- Χωρική διακριτοποίηση (κεντρικές διαφορές)



$$\begin{aligned} \text{div}(g\nabla u)|_{ij} \approx & \frac{1}{h_x^2} \{ g_{i+\frac{1}{2},j} (u_{i+1,j} - u_{ij}) - g_{i-\frac{1}{2},j} (u_{ij} - u_{i-1,j}) \} \\ & + \frac{1}{h_y^2} \{ g_{i,j+\frac{1}{2}} (u_{i,j+1} - u_{ij}) - g_{i,j-\frac{1}{2}} (u_{ij} - u_{i,j-1}) \} \end{aligned}$$

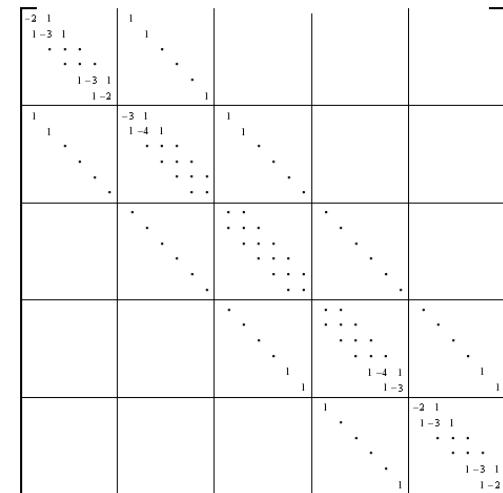
- Μορφή γινομένου πίνακα-διανύσματος

$$\text{div}(g\nabla u) \approx Au$$

- Ημι-διακριτοποιημένη μορφή:

$$\frac{\partial u}{\partial t} = Au + F(u)$$

$$A =$$



Χρονική ολοκλήρωση

■ Εμπρόσθιο σχήμα

$$(u^{n+1} - u^n)/\tau = Au^n + F(u^n) \quad t = \tau, 2\tau, \dots, n\tau, \dots$$

□ Απλό αλλά περιοριστικό: $\tau \leq h^2/4$

■ Πλήρως οπίσθιο σχήμα

$$(u^{n+1} - u^n)/\tau = Au^{n+1} + F(u^{n+1})$$

□ Ευσταθές αλλά δυσεπίλυτο (ισχυρά μή-γραμμικό)

■ Υβριδικό σχήμα (IMEX)

$$(u^{n+1} - u^n)/\tau = Au^{n+1} + F(u^n)$$

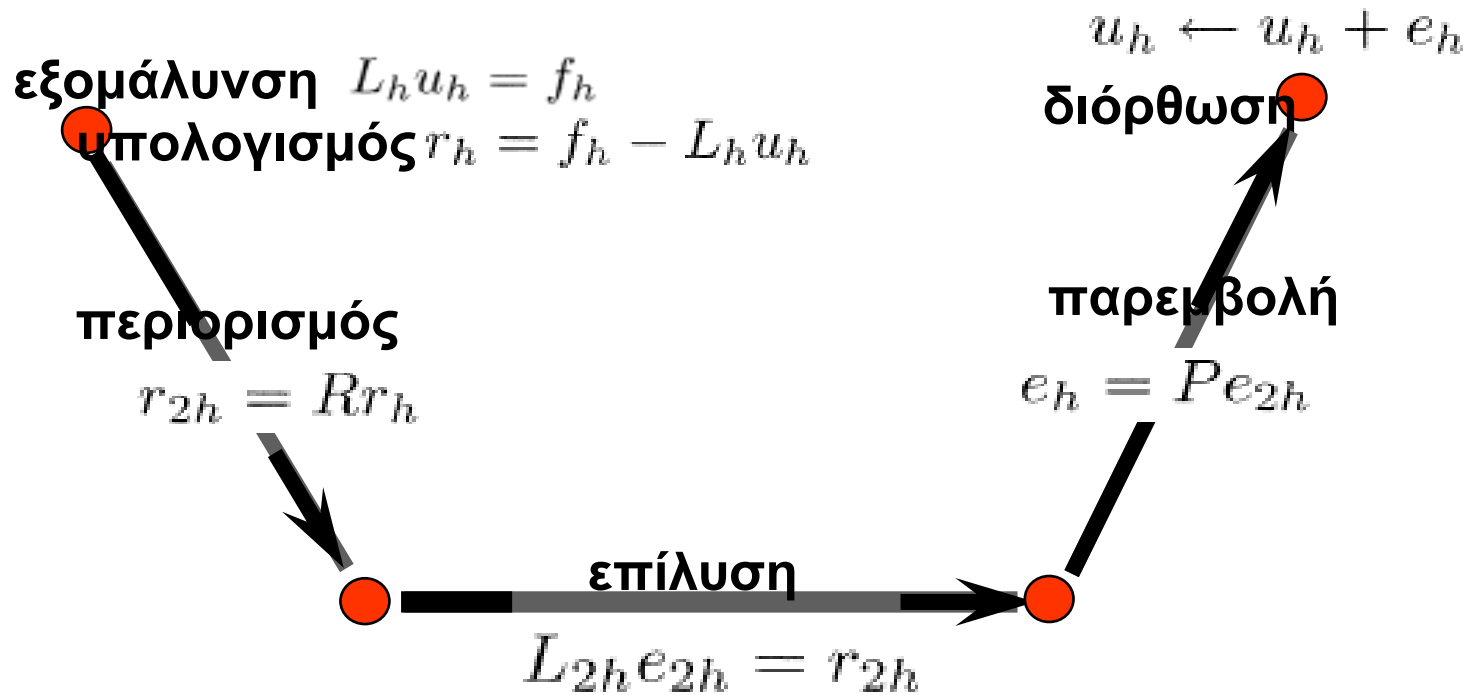
□ Ευσταθές, αλλά επίλυση γραμμικού συστήματος/βήμα:

$$\left[\frac{1}{\tau}I - A \right] u^{n+1} = \frac{1}{\tau}u^n + F(u^n)$$

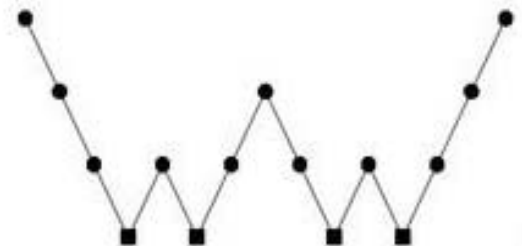
πίνακας συστήματος
 $L \equiv (1/\tau)I - A$

Επίλυση με πολυπλεγματικές τεχνικές!

Πολυπλεγματοί αλγόριθμοι



- Αναδρομική επέκταση σε πολλά επίπεδα (V/W κύκλοι)
- Πολυπλεγματοί συστατικά:
 - Εξομαλυντής
 - Πίνακες περιορισμού R και παρεμβολής P
 - Πρόβλημα στο αδρό επίπεδο L_{2h}



Υπολογιστικό κόστος

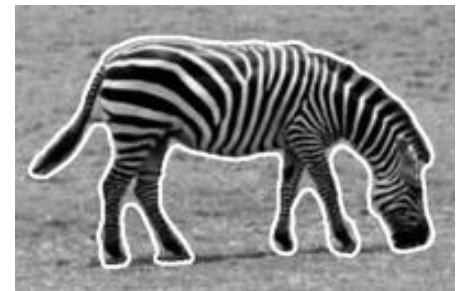
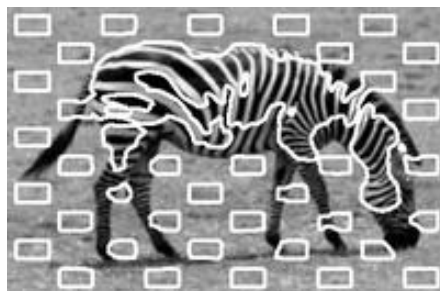
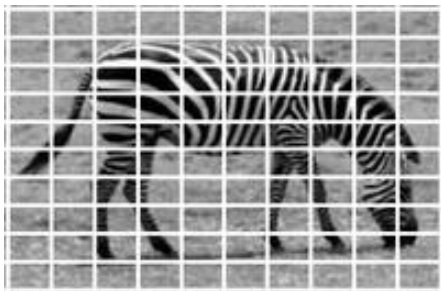
- Ενδεικτικό υπολογιστικό κόστος (ανά χρονικό βήμα)

Λειτουργία		Αλγόριθμος	Χρόνος (msec)	
			Αρχικοποίηση	Ανά βήμα
Σχήμα	2-Δ IMEX	Πολυπλεγματικός (Γεωμ. - MDMG)	35±1	23±1
		Πολυπλεγματικός (Αλγεβ. - AMG)	300±20	40±5
	AOS	Τριδιαγώνιος Επιλύτης (TRI)	13±1	5±1
Μετασχ. Απόστασης		Διαχωρίσιμος (SDT)	-	27±1
		Ταχεία Προέλαση (FMDT)	-	38±2

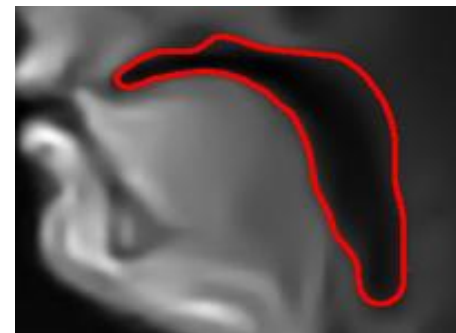
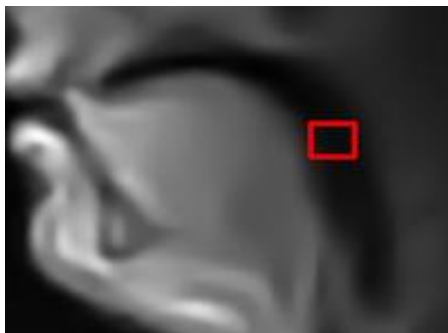
- Βιβλιοθήκη GAC++. Πηγαίος κώδικας υπό GPL άδεια
<http://cvsp.cs.ntua.gr/software/GAC++>
- Ευρεία χρήση στο εργαστήριό μας και αλλού

Γεωμετρικά ενεργά περιγράμματα: Πειράματα κατάτμησης

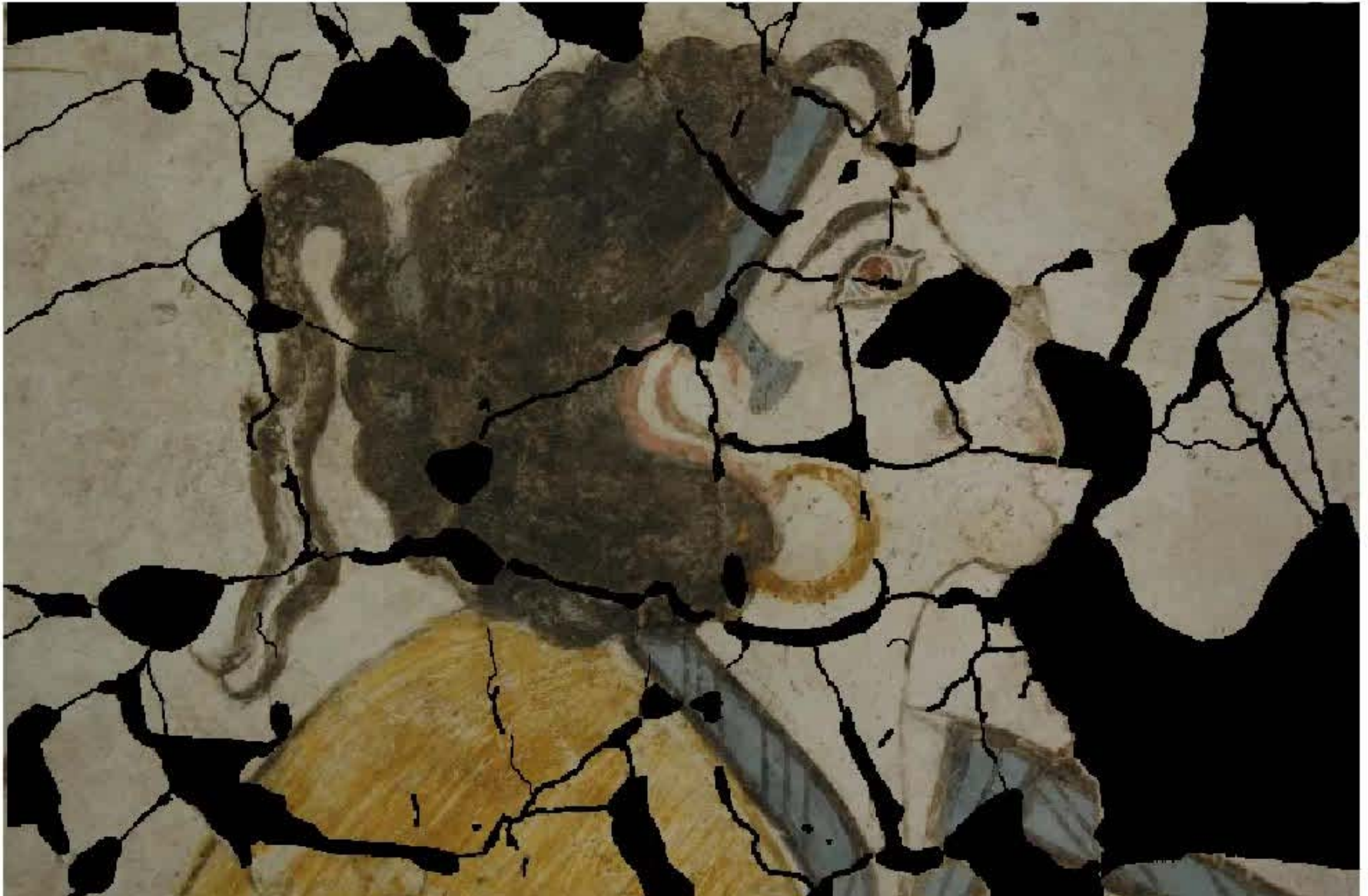
■ Κατάτμηση με βάση ομοιογένεια υφής



■ Κατάτμηση με βάση ομοιογένεια φωτεινότητας



Παράδειγμα ενδοσυμπλήρωσης με ΜΔΕ διάχυσης



Λεπτομέρειες Αποτελέσματος Ενδοσυμπλήρωσης

- Επαναληπτική διαδικασία, ευνοεί συνέχιση ακμών



αρχικό



ενδιάμεσο

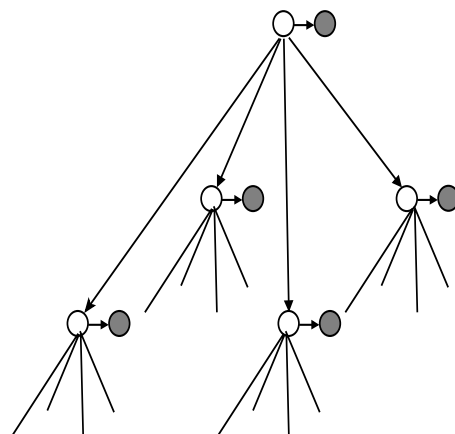


τελικό

- Πιθανή δημιουργία ψευδέργων (artifacts)

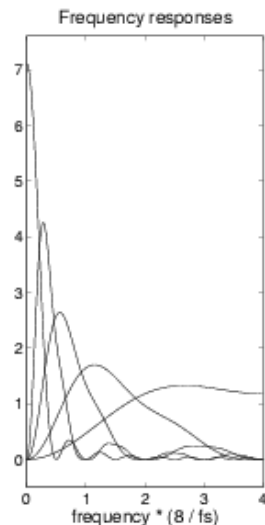
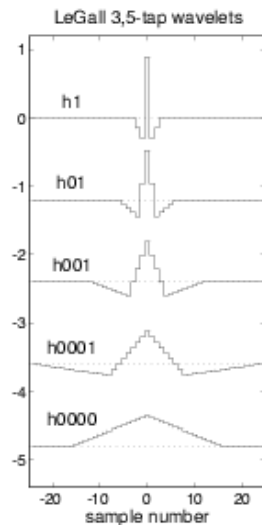
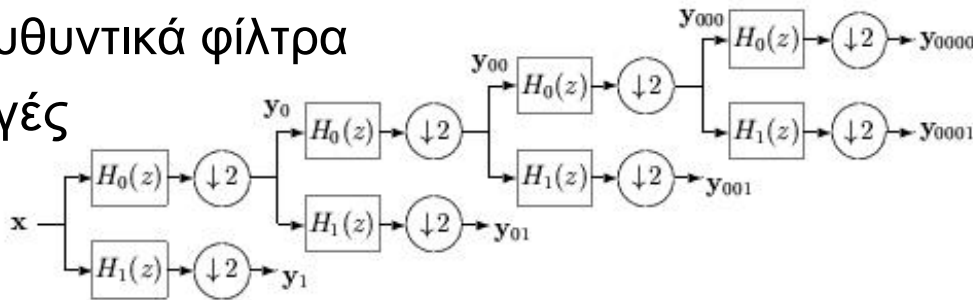


(Μέρος IV) Ενδοσυμπλήρωση στο πεδίο των κυματιδίων (wavelets)

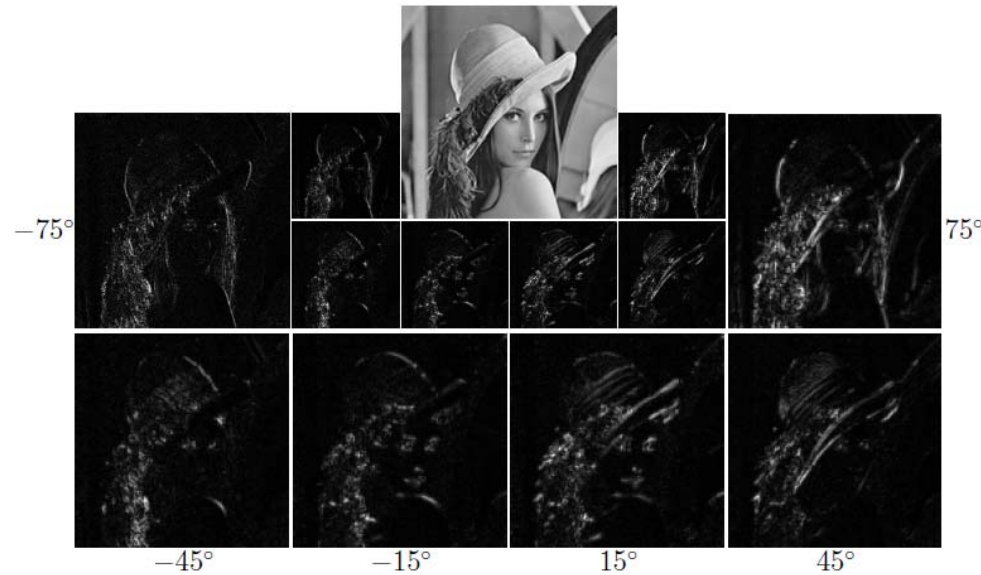
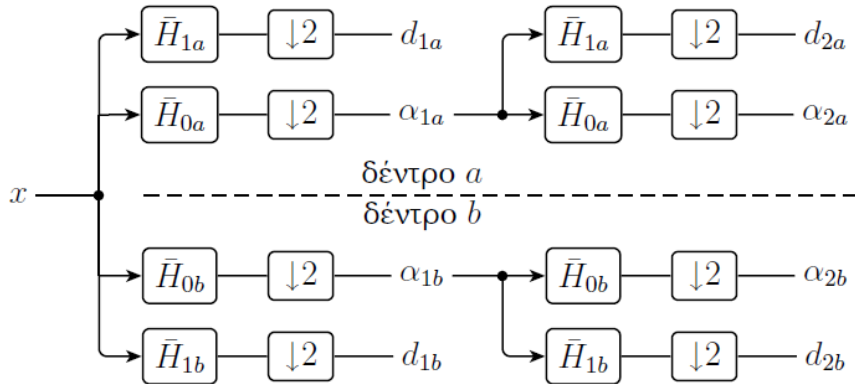
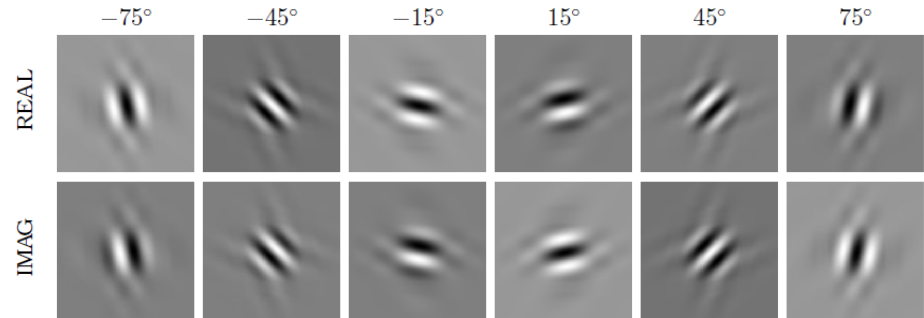
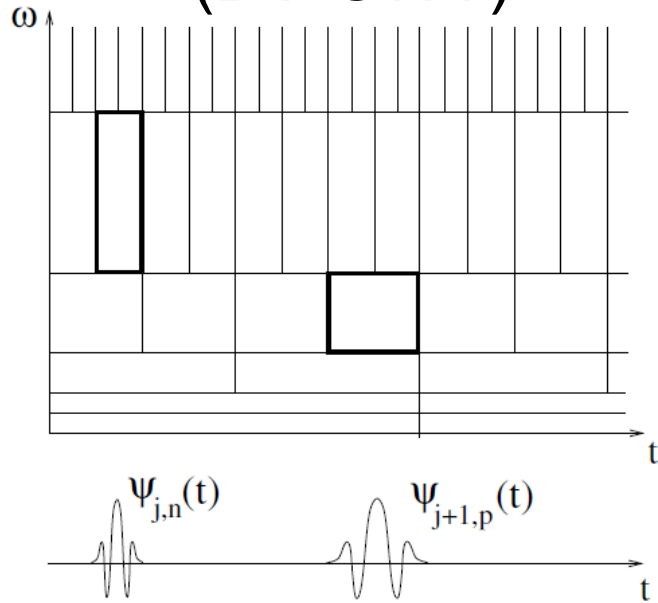


Ανάλυση και επεξεργασία εικόνων με κυματίδια

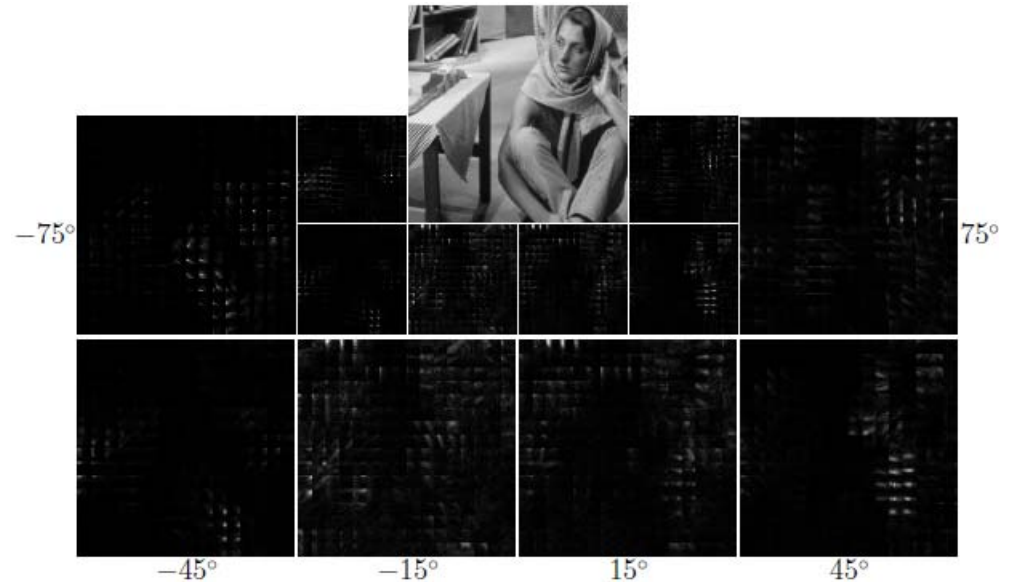
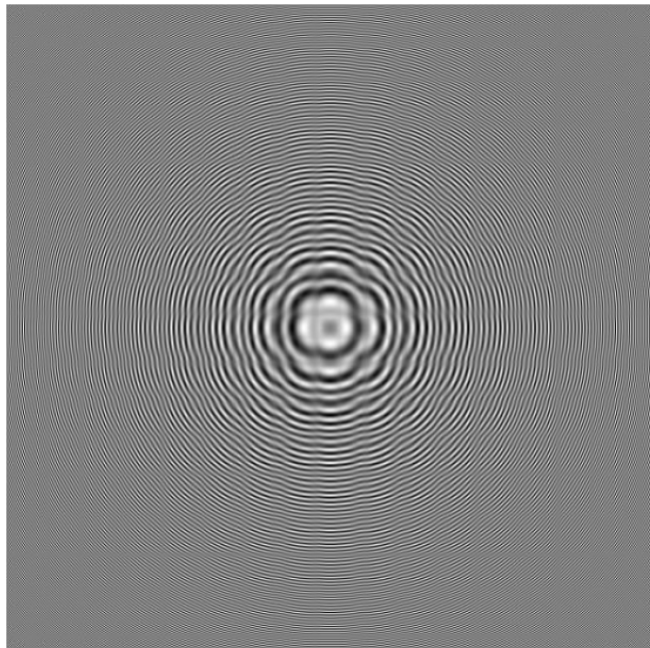
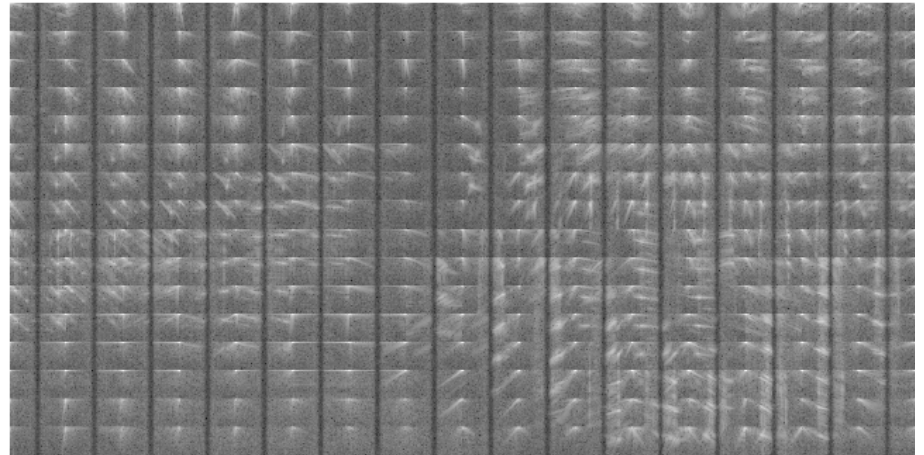
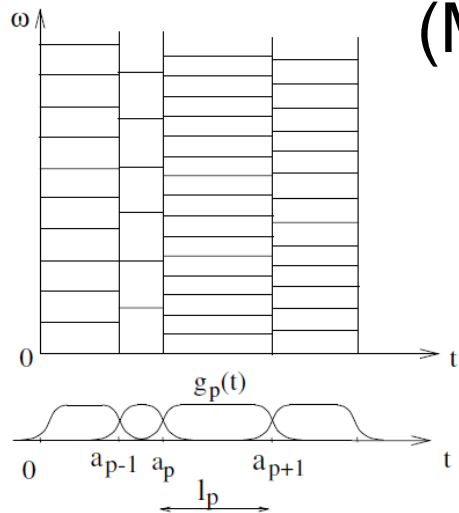
- Κυματίδια: πολυκλιμακωτή ανάλυση εικόνων
 - Γραμμικά, ζωνοπερατά, κατευθυντικά φίλτρα
- Σημαντική επιτυχία σε εφαρμογές
 - Κωδικοποίηση εικόνων
 - Αποθορυβοποίηση
 - Συμπιεσμένη αίσθηση (compressed sensing)
- Βάση για αποδοτικούς αλγορίθμους
 - $O(N)$ ευθύς/αντίστροφος διακριτός μετασχ. κυματιδίων



Δίδεντρος Μιγαδικός Κυματιδιακός Μετασχηματισμός (DT-CWT)



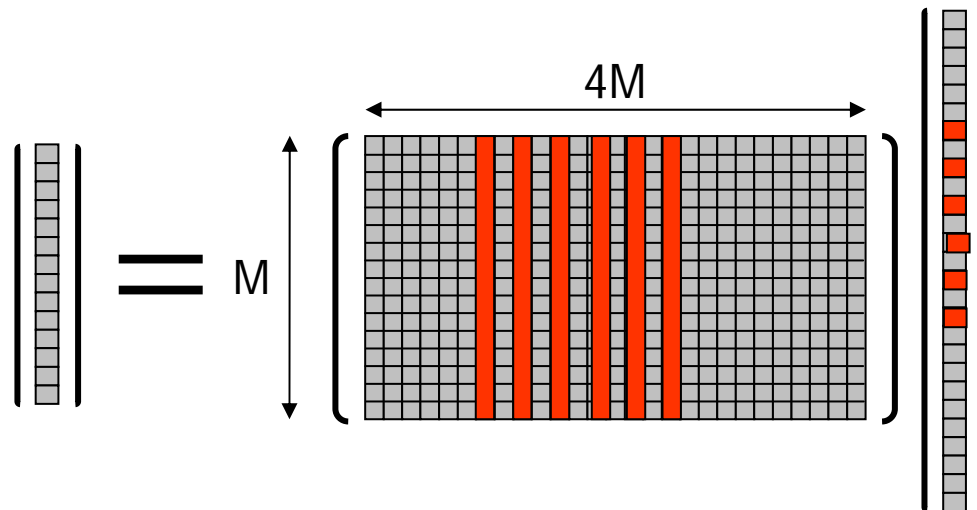
Διαμορφωμένος Μιγαδικός Αλληλεπικαλυπτόμενος Μετ/σμός (MCLT)



Αναπαράσταση εικόνας με μιγαδικά κυματίδια

Ιδιότητες (DT-CWT και MCLT)

- ❑ Κατευθυντικότητα (6 μιγαδικές υποζώνες / κλίμακα)
- ❑ Συμμετρία μετατόπισης
- ❑ Υπερ-πλήρη (4-φορές) πλαίσια
- ❑ Ανάλυση χώρου-κλίμακας έναντι χώρου-συχνότητας



$$x = F_s y$$

x τιμές εικόνας
 y συντ. κυματιδίων
 F_s IDWT

Κυματίδια και στατιστική ενδοσυμπλήρωση

- Αντίστροφο πρόβλημα
- Πιθανοτική προσέγγιση στο πεδίο κυματιδίων:

1. Μετρήσεις (γνωστά σημεία) $P(x_o | y)$

2. Πρότερο μοντέλο κυματιδίων $P(y)$

Κανόνας
Bayes

3. Συμπερασμός για κυματίδια: $P(y | x_o), \hat{y}$

IDWT

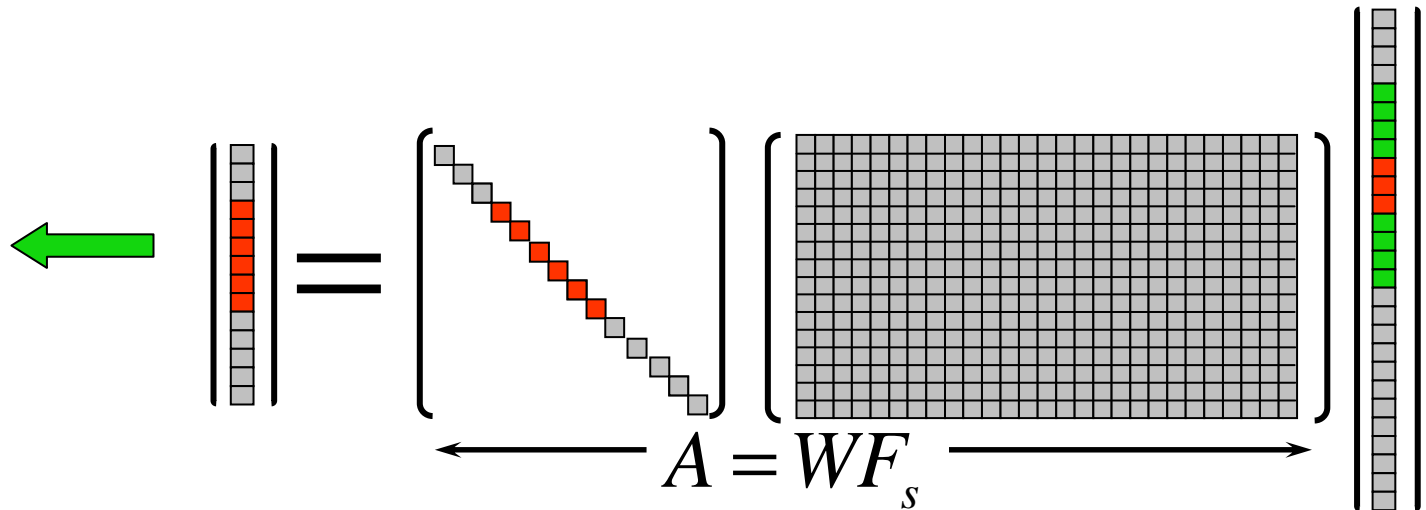
4. Ενδοσυμπλήρωση $\hat{x} = F_s \hat{y}$

Μοντέλο μέτρησης

- Παρατηρημένες/επικαλυμμένες τιμές

$$x = \{x_o, x_h\}$$

$$x_o = WF_s y + n = Ay + n \Rightarrow P(x_o | y) = N(x_o; Ay, \Sigma_n)$$



1. Μετρήσεις (γνωστά σημεία) $P(x_o | y)$

2. Πρότερο μοντέλο κυματιδίων $P(y)$

Κανόνας
Bayes

3. Συμπερασμός για κυματίδια $P(y | x_o), \hat{y}$

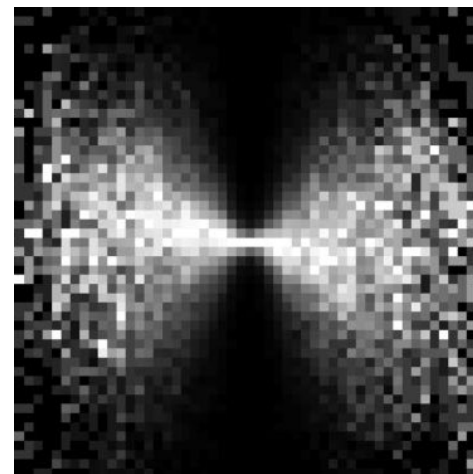
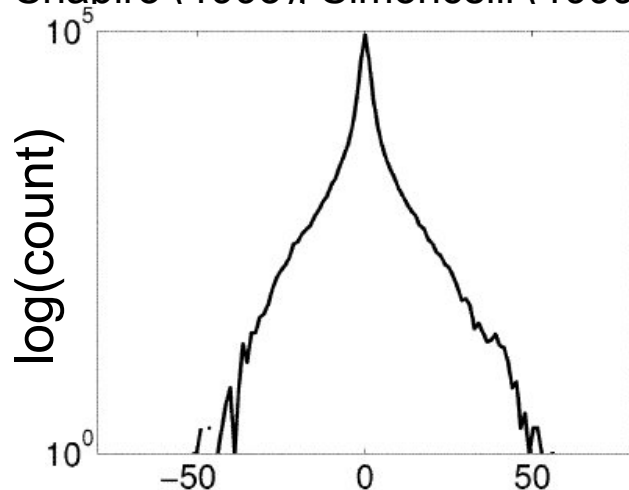
IDWT

4. Ενδοσυμπλήρωση $\hat{x} = F_s \hat{y}$

Στατιστική εικόνων στο πεδίο κυματιδίων

- Ιδιότητα αποσυσχέτισης: $E\{y_i y_j\} \approx \delta_{ij}$
- Μη-γκουσιανές στατιστικές ιδιότητες
 - Λεπτοκυρτωτικά περιθώρια ιστογράμματα
- Αλληλεξαρτήσεις υψηλής τάξης σε γειτονιές χώρου/κλίμακας
 - Από-κοινού ιστογράμματα «πεταλούδας»

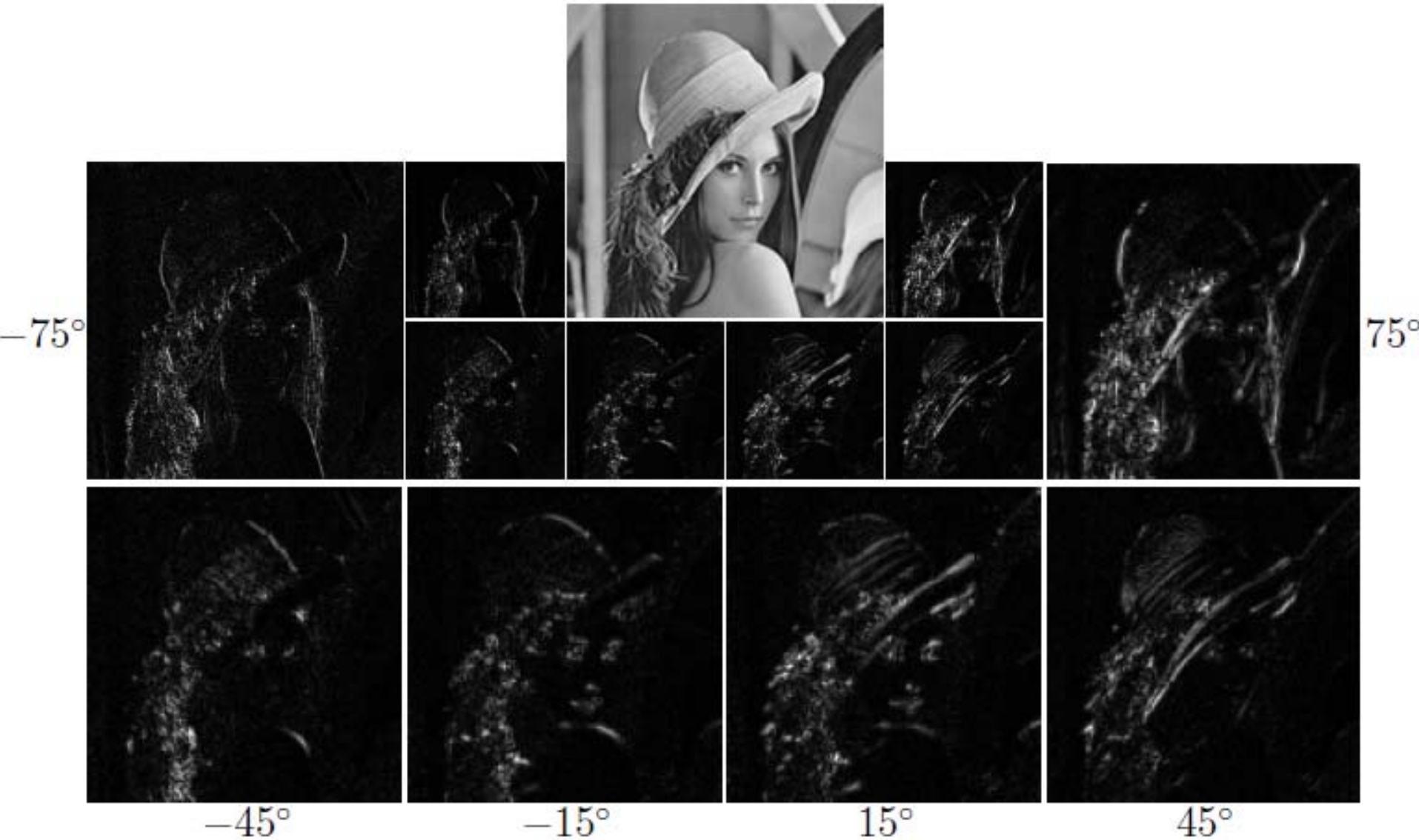
Shapiro (1993). Simoncelli (1999), Portilla et al. (2003), ...



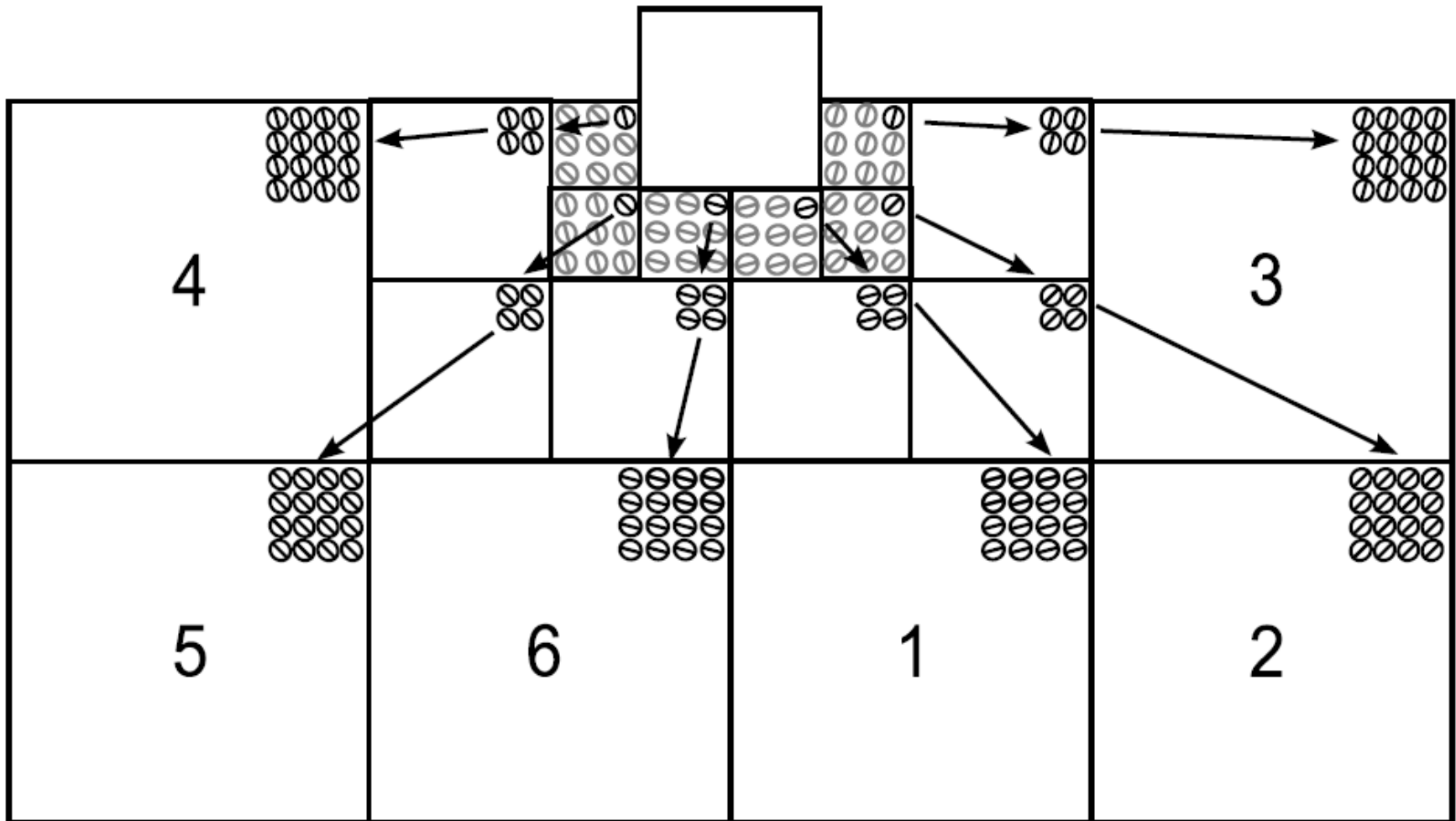
Μοντελοποίηση εικόνων στο πεδίο κυματιδίων

- Λεπτοκυρτωτικές (αραιές) κατανομές συντελεστών
 - Laplacian (L_1) / Student-t κατανομές
 - Μείγμα κατανομών Gaussians
- Αλληλεξαρτήσεις μεταξύ κλιμάκων
 - Αλληλουχίες από Gaussian scale mixtures
 - Κρυφό δένδρο Markov
- Σύγκριση με βιβλιογραφία που συνδυάζει κυματίδια & ενδοσυμπλήρωση:
 - Προηγούμενες εργασίες: κυρτωτικές κατανομές
 - Συνεισφορά μας: + **μοντελοποίηση εξαρτήσεων**

Δομή κρυφού δέντρου Markov στο πεδίο μιγαδικών κυματιδίων

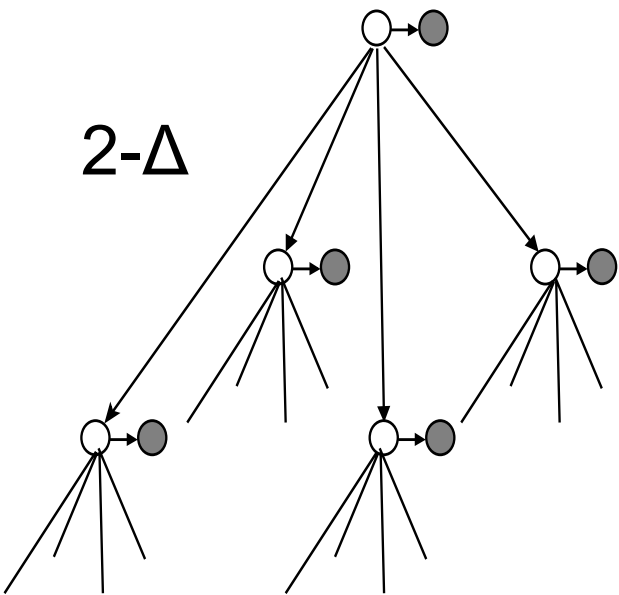


Δομή κρυφού δέντρου Markov στο πεδίο μιγαδικών κυματιδίων



Κρυφό δέντρο Markov (HMT) στο πεδίο κυματιδίων

2- Δ



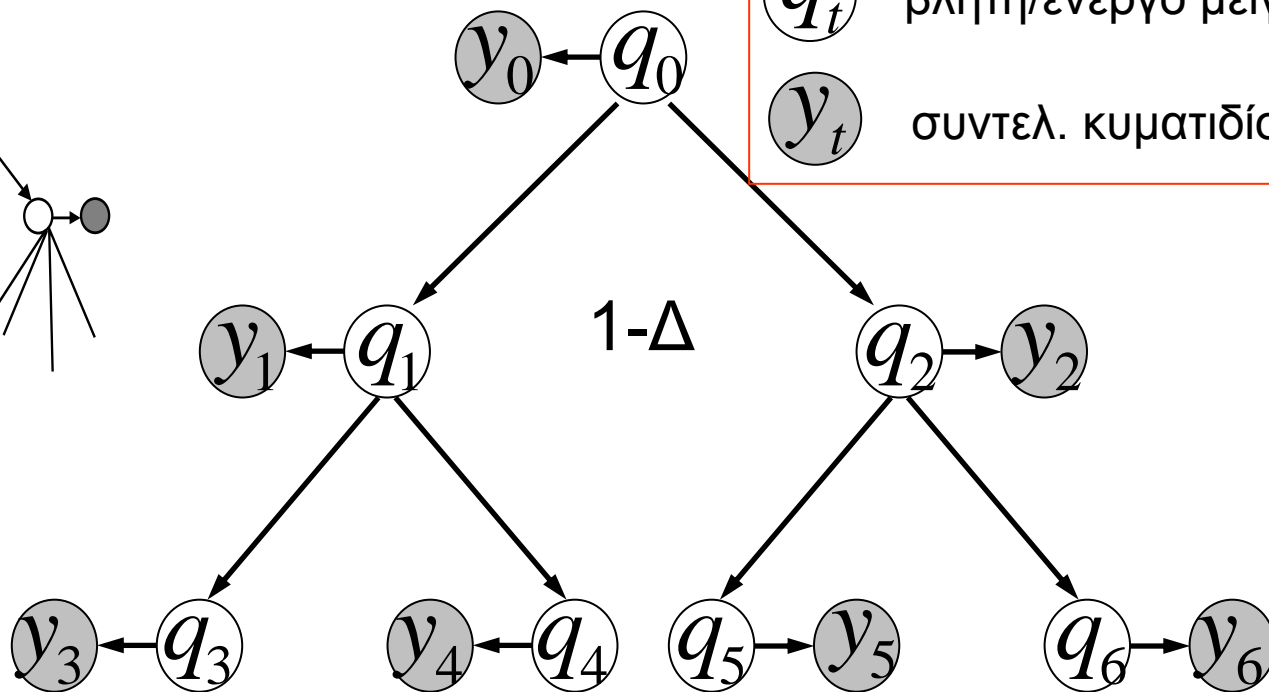
q_t

διακριτή κρυφή μεταβλητή/ενεργό μείγμα

y_t

συντελ. κυματιδίου

1- Δ



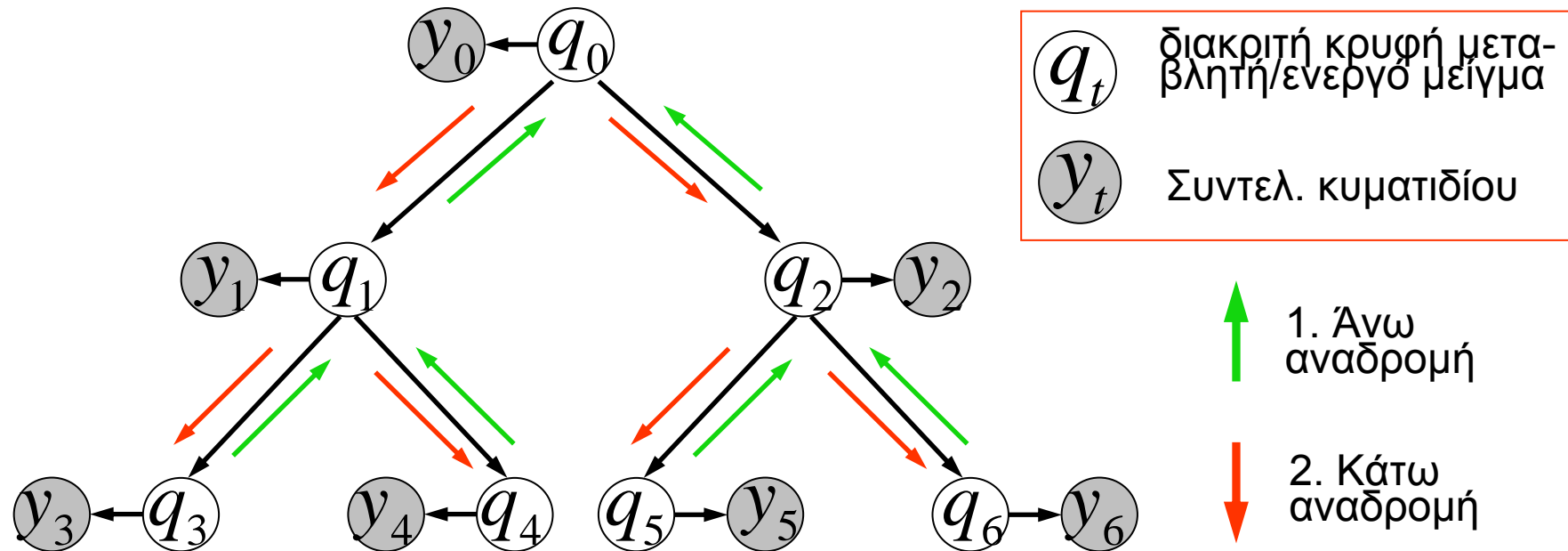
$$P(q) = \pi_{q_0} \prod_{t=1}^T \alpha_{q_{p_t} q_t}$$

θ παράμετροι μοντέλου
 $= \{ \pi, \alpha_{ji}, \mu, \Sigma \}$

p_t γονιός κόμβου t
 $T+1$ πλήθος κόμβων
 Q κρυφές καταστ.

$$P(y | q) = N(y; \mu_q, \Sigma_q) = \prod_{t=0}^T N(y_t; \mu_{q_t}, \Sigma_{q_t})$$

Κρυφά δέντρα Markov: Βασικές λειτουργίες

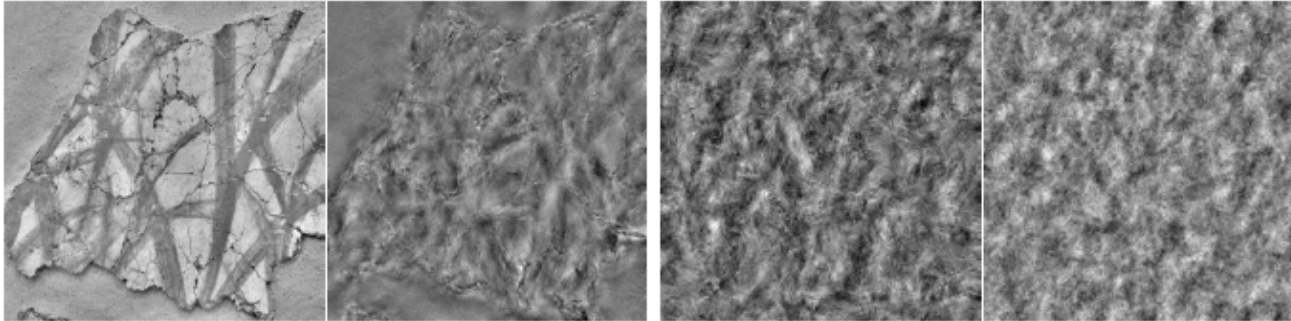


Αποδοτικοί αναδρομικοί ως προς την κλίμακα αλγόριθμοι:

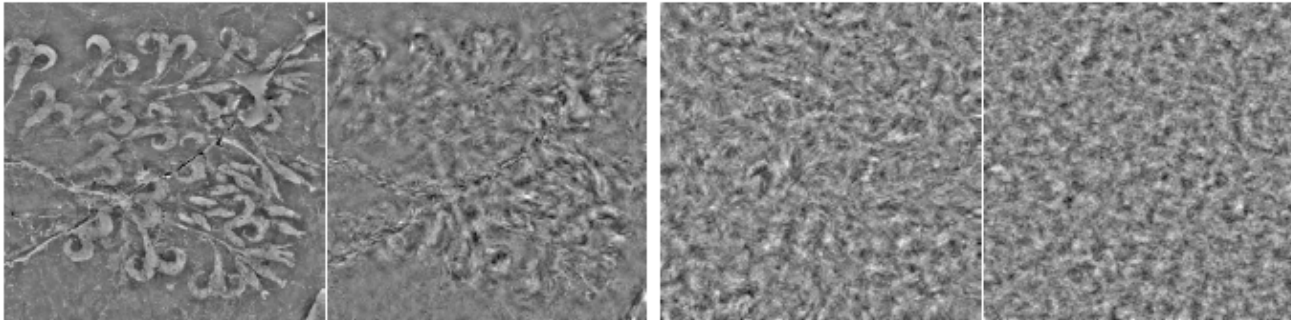
- Αποτίμηση πιθανοφάνειας (άνω) $P(y)$
- Συμπερασμός (άνω-κάτω) $P(q_t | y)$
- Viterbi πιθανότερο μονοπάτι (άνω-κάτω) $\hat{q} = \arg \max P(q | y)$
- Παράμετροι μέγ. πιθανοφάνειας με EM $\hat{\vartheta} = \arg \max P(y; \vartheta)$
- Υπό-συνθήκη δειγματολ. (άνω-κάτω) $q_s \sim P(q | y)$

Πειράματα σύνθεσης υφής με το HMT μοντέλο

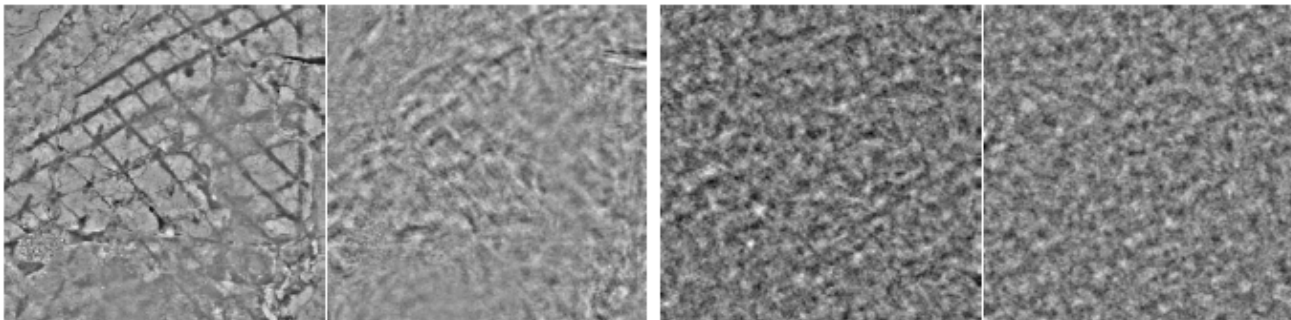
καλάμια - λεπτομέρεια Σχ. 2.5(α)



κρίνα - λεπτομέρεια Σχ. 2.5(γ)



φόρεμα - λεπτομέρεια Σχ. 2.5(γ)



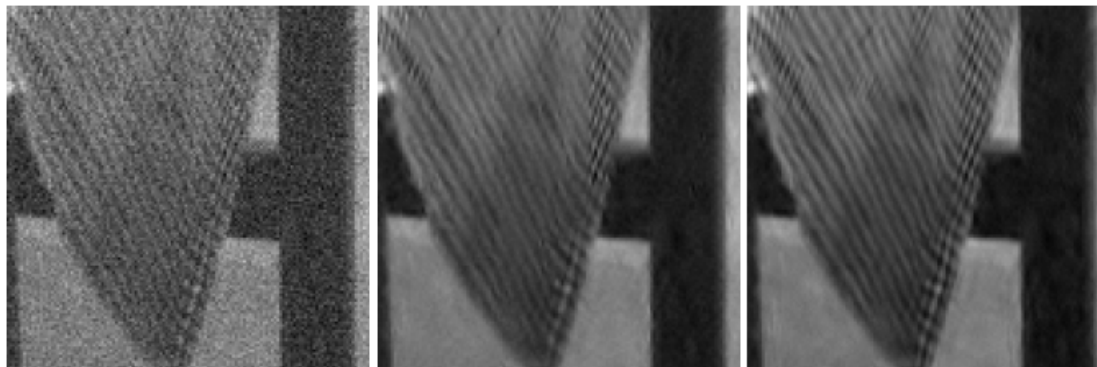
αρχική

ασύμφωνη
αρχική

σύνθεση
HMT

σύνθεση
HIT

Πειράματα αποθορυβοποίησης με το HMT μοντέλο



θορυβώδης

HMT-DTCWT

HMT-MCLT



θορυβώδης

HMT-DTCWT

HIT-DTCWT

Συστηματική αξιολόγηση (7 εικόνες αναφοράς, 4 επίπεδα θορύβου):

(1) HMT κέρδος +0.4 dB (μέσος όρος)

(2) MCLT κέρδος ως +0.7 dB (ορισμένες εικόνες)

(3) Επιδόσεις κοντά στο state-of-art

- **Συνεισφορά:** EM αλγόριθμος για εκπαίδευση HMT μοντέλου από θορυβώδη δεδομένα

1. Μετρήσεις (γνωστά σημεία) $P(x_o | y)$

2. Πρότερο μοντέλο κυματιδίων $P(y)$

Κανόνας
Bayes

3. Συμπερασμός για κυματίδια $P(y | x_o), \hat{y}$

IDWT

4. Ενδοσυμπλήρωση $\hat{x} = F_s \hat{y}$

Συμπερασμός για Κυματιδιακούς Συντελεστές

- Ακριβής συμπερασμός αδύνατος

$$P(y | x_o) \propto P(x_o | y)P(y) =$$

$$= P(x_o | y) \sum_{\text{all paths } q} P(y | q)P(q)$$

άθροιση όλων των μονοπατιών $O(Q^T)$

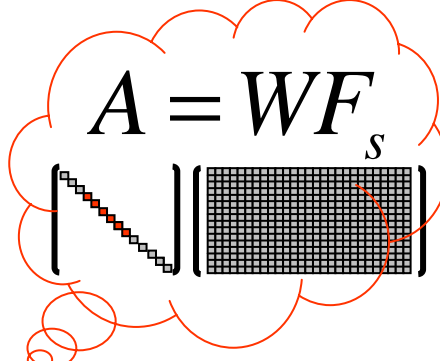
- Όμως, δοσμένης της κρυφής κατάστασης, κυματιδιακή εκτίμηση εφικτή (πρόβλημα ελαχίστων τετραγώνων):

$$P(y | x_o, q) \propto P(x_o | y)P(y | q) =$$

$$= N(y; \mu_{y|q}, \Sigma_{y|q})$$

$$\mu_{y|q} = \Sigma_{y|q} (\Sigma_q^{-1} \mu_q + A^T \Sigma_n^{-1} x_o)$$

$$\Sigma_{y|q}^{-1} = \Sigma_q^{-1} + A^T \Sigma_n^{-1} A$$

$$A = W F_s$$


Αλγόριθμος I: Εναλλασσόμενη βελτιστοποίηση

1. Αρχική εκτίμηση y_0

2. Δοσμένης εκτίμησης y_{i-1} για συντελ. κυματιδίων, εύρεση του πιο πιθανού (Viterbi) μονοπατιού

$$q_i = \arg \max P(q | y_{i-1})$$

3. Δοσμένης εκτίμησης q_i για το μονοπάτι, ενημέρωση συντελεστών κυματιδίων

$$y_i = \arg \max P(y | q_i, x_o) = \mu_{y|q_i}$$

4. Έξοδος
 $\hat{x} = F_s y_\infty$

□ Block ICM τεχνική

□ Μονοτονική σύγκλιση σε τοπικό μέγιστο $P(y, q | x_o)$

$$\mu_{y|q_i} = \sum_{y|q_i} (\Sigma_{q_i}^{-1} \mu_{q_i} + A^T \Sigma_n^{-1} x_o)$$

$$\Sigma_{y|q_i}^{-1} = \Sigma_{q_i}^{-1} + A^T \Sigma_n^{-1} A$$

Αλγόριθμος II: Δειγματοληψία Gibbs

1. Αρχική εκτίμηση y_0

2. Δοσμένου δείγματος κυματιδίων y_{i-1} , λήψη μπλοκ δείγματος μονοπατιού

$$q_i \sim P(q | y_{i-1})$$

3. Δοσμένου μονοπατιού q_i , λήψη μπλοκ δείγματος συντελεστών κυματιδίων

$$y_i \sim P(y | q_i, x_o) = y_s$$

- ❑ Μπλοκ Gibbs δειγματοληψία της $P(y, q | x_o)$
- ❑ Πλήρης διερεύνηση της ύστερης κατανομής

$$y_s = \sum_{y|q_i} (\Sigma_{q_i}^{-1} y_{u,s} + A^T \Sigma_n^{-1} x_{o,s})$$

$$\Sigma_{y|q_i}^{-1} = \Sigma_{q_i}^{-1} + A^T \Sigma_n^{-1} A$$

Πειράματα ενδοσυμπλήρωσης

καλυμμένη



ΜΔΕ Laplace



PSNR
20.6 dB

Προτ. Μοντέλο
(Alt. Opt.)



PSNR
21.9 dB

Προτ. Μοντέλο
(Δειγμ. Gibbs)



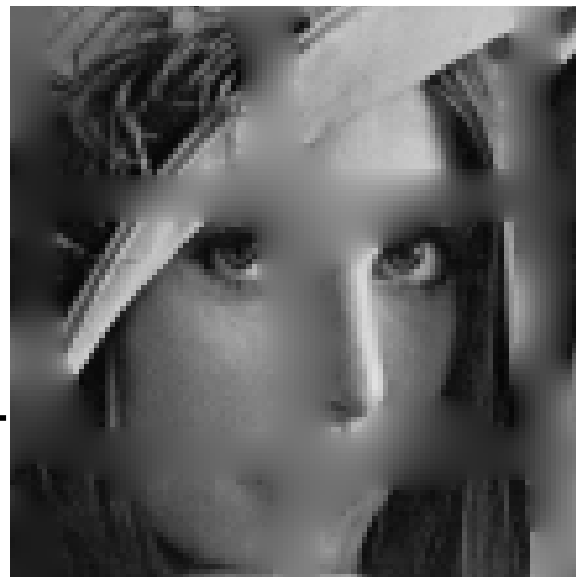
PSNR
22.0 dB

Πειράματα ενδοσυμπλήρωσης (μεγέθυνση)

occluded

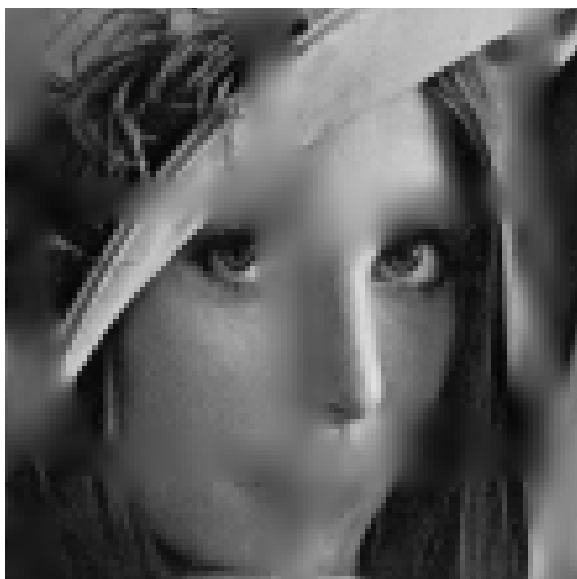


Laplace PDE



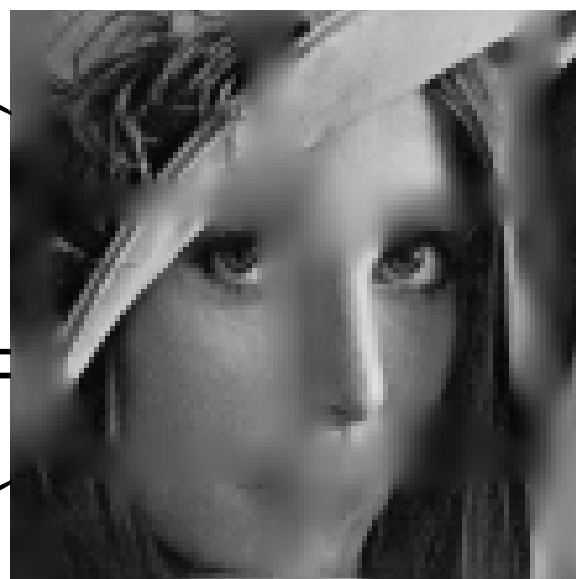
PSNR
20.6 dB

Προτ. Μοντέλο
(Alt. Opt.)



PSNR
21.9 dB

Προτ. Μοντέλο
(Δειγμ. Gibbs)



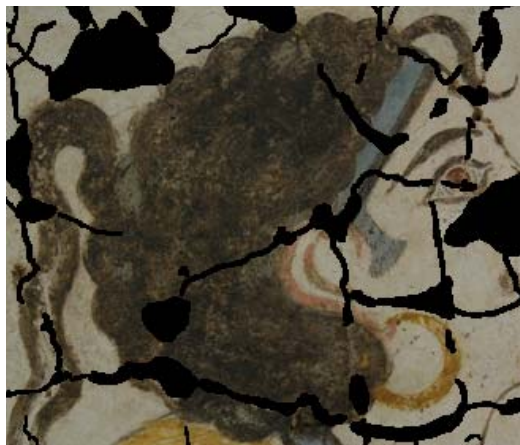
PSNR
22.0 dB

Αποτελέσματα σε Τοιχογραφίες Αιρωτηρίου



Λεπτομέρειες Αποτελέσματος Ενδοσυμπλήρωσης

- Ικανοποιητική συμπλήρωση ομαλών περιοχών/μικρών κενών



επικαλυμμένη

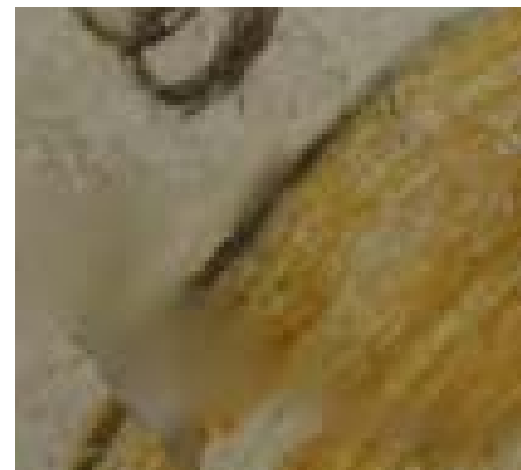


αρχική συνθήκη

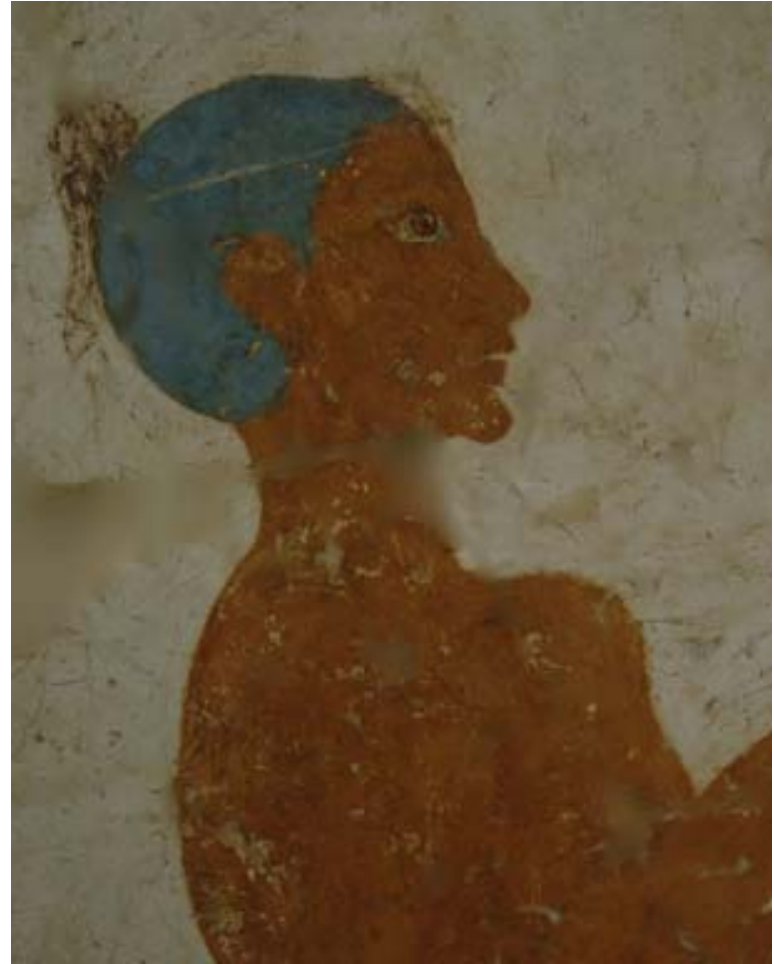
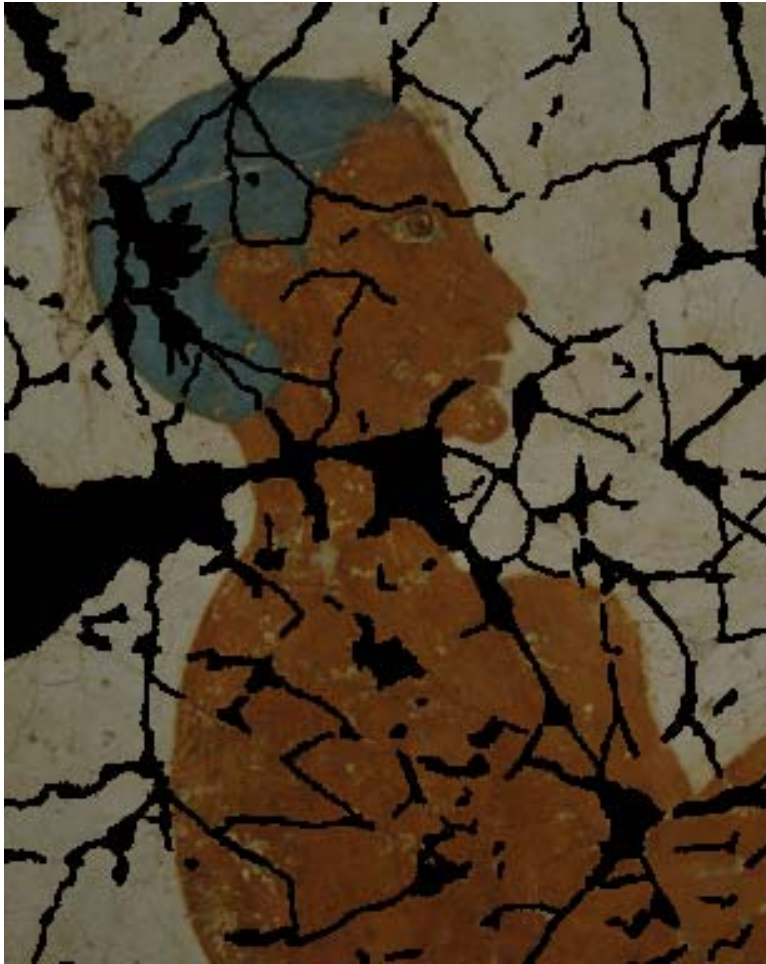


λύση στη σύγκλιση

- Χωρίς ψεύδεργα, αλλά αδυναμία συνένωση ακμών

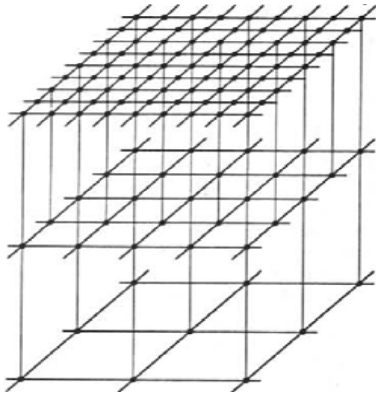


Ακολουθία Ανδρών



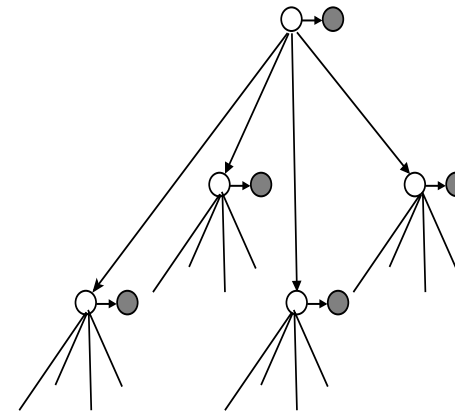
Σύνοψη Τεχνικών ΜΔΕ και Κυματιδίων

ΜΔΕ



- Μοντέλα φυσικής
- Χωρικό πεδίο (γεωμετρία)
- Ντετερμινιστική επίλυση

έναντι



Κυματίδια

- Στατιστικά μοντέλα
- Πεδίο κυματιδίων (στατιστική)
- Στοχαστική επίλυση

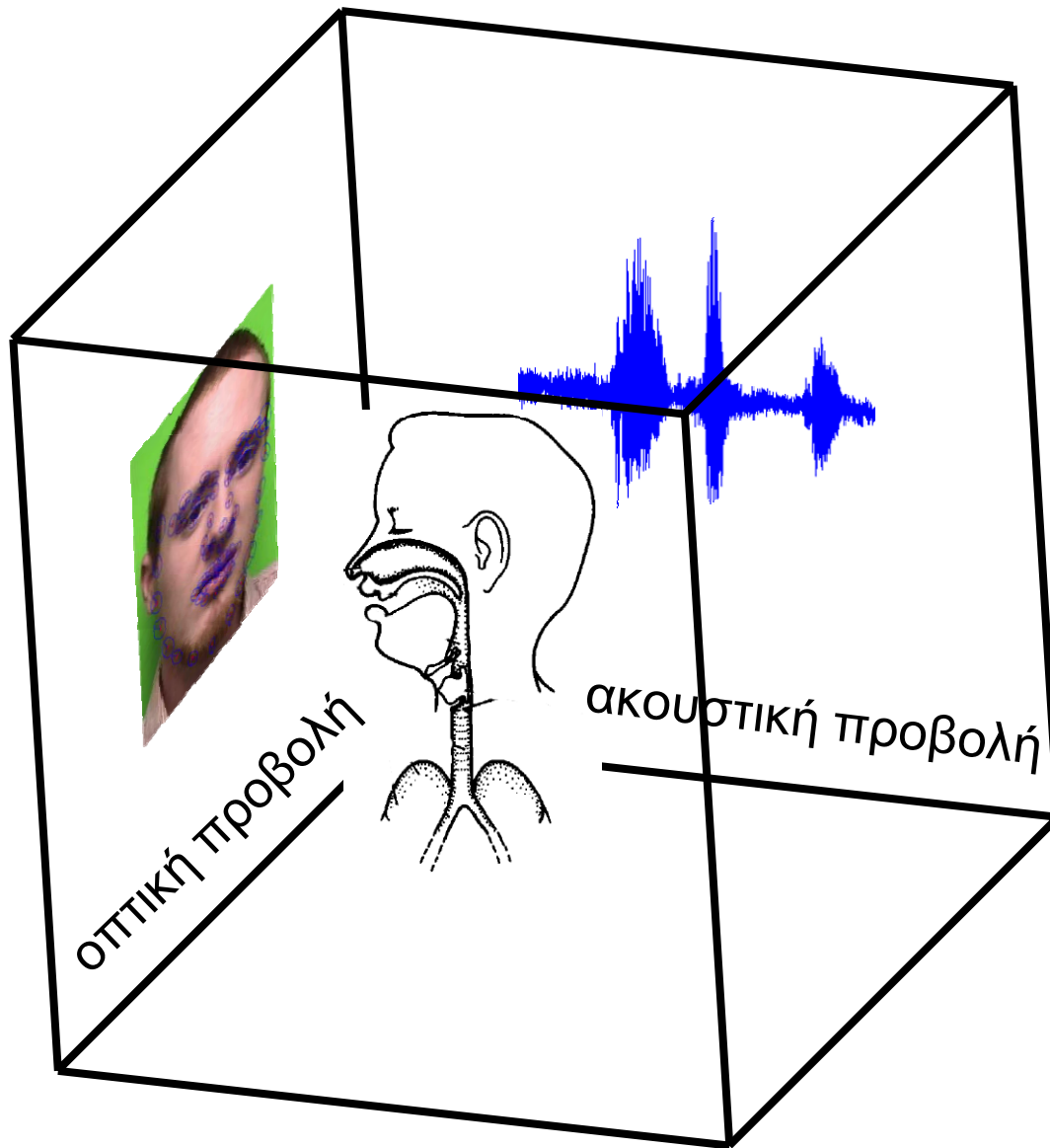
- Πολυκλιμακωτές αναπαραστάσεις
- Αποδοτικοί $O(N)$ αλγόριθμοι
- Υβριδική προσέγγιση για το πρόβλημα της ενδοσυμπλήρωσης

Περιοχές μελλοντικής έρευνας

- Αυτόματη ανίχνευση ελλειπουσών περιοχών στις τοιχογραφίες
 - Πιθανόν να επιβοηθάται από 3-Δ πληροφορία
 - Επιτρέπει εφαρμογή σε μεγαλύτερη κλίμακα στις τοιχογραφίες Ακρωτηρίου
- Πρακτική αξιοποίηση στο Ακρωτήρι
 - Εργαλείο μελέτης και ανάδειξης των τοιχογραφιών
 - Αξιοποίηση στο πρόβλημα της εύρεσης επαφών
- Ενδοσυμπλήρωση με τεχνικές κυματιδίων
 - Ανάπτυξη & ενσωμάτωση μοντέλου μιγαδικής φάσης
 - Ενσωμάτωση ιδεών από μέσο επίπεδο όρασης
 - Αντικειμενοστρεφή ιεραρχικά αναγεννητικά μοντέλα εικόνων

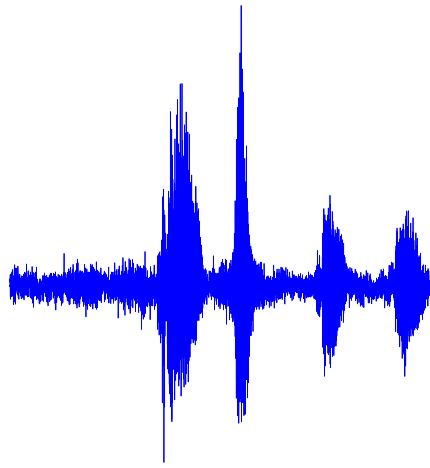
Αναγνώριση Ομιλίας από Ακουστική και
Οπτική Πληροφορία

Ομιλία: Πολύπλευρο φαινόμενο

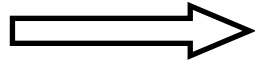


CVSP

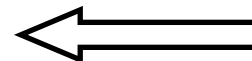
Αναγνώριση Ομιλίας από Ήχο και Εικόνα



Ήχος



one two
three
six four
five



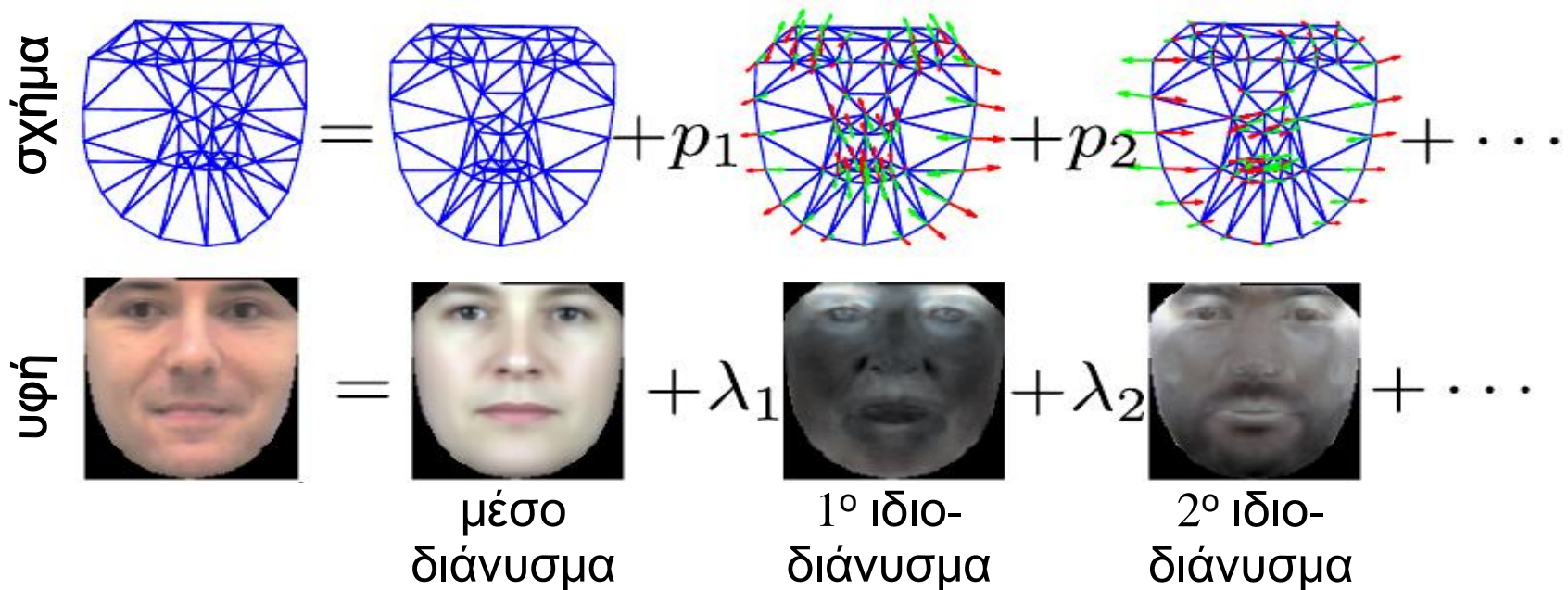
Εικόνα

Κύρια σημεία:

- Εξαγωγή οπτικών χαρακτηριστικών από το πρόσωπο ομιλητή
- Σύμμιξη οπτικής και ακουστικής πληροφορίας

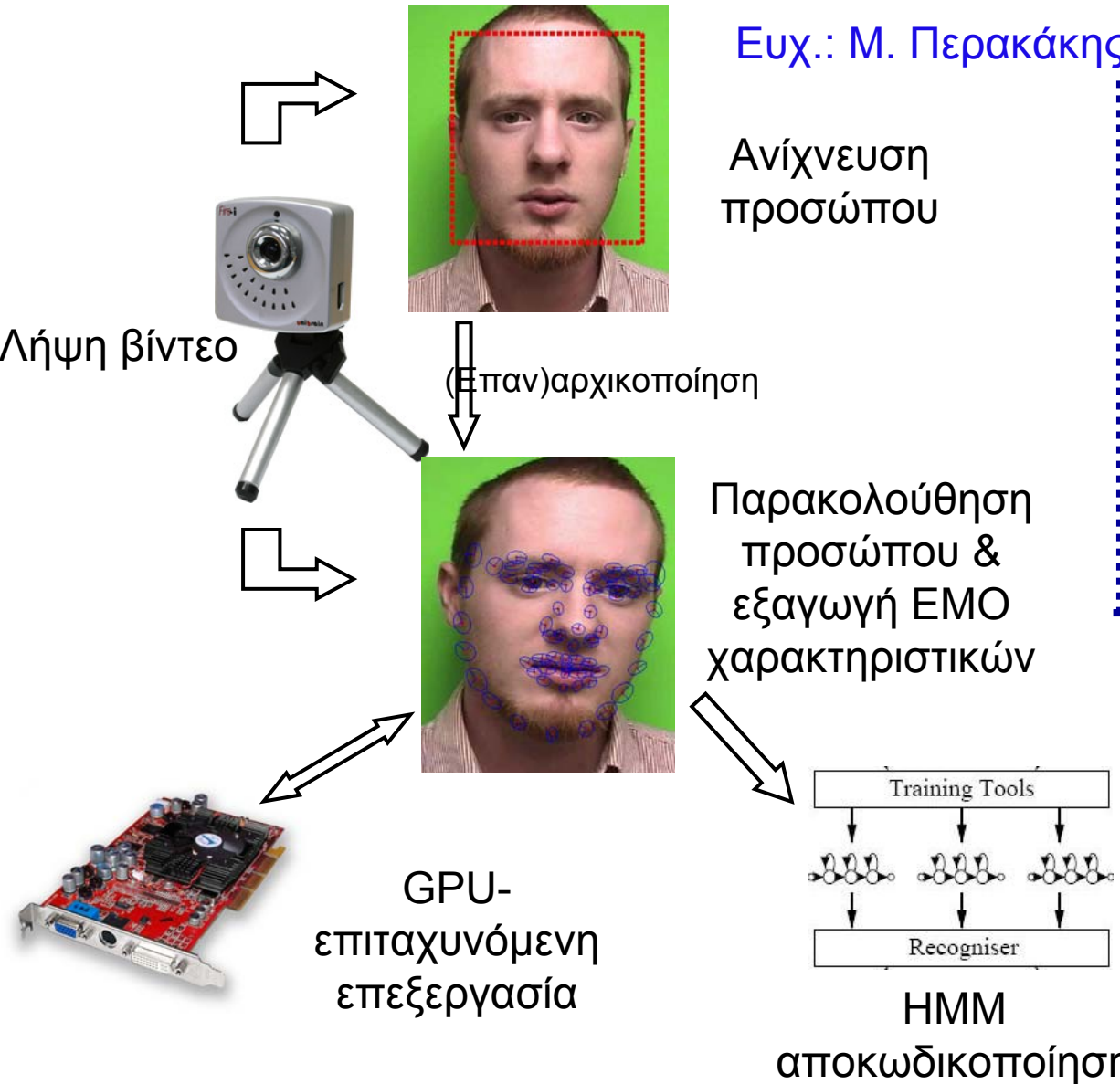
Εξαγωγή Οπτικής Πληροφορίας

- Σχήμα και υφή του προσώπου επιβιβάζονται το διάβασμα των χειλιών
- Μοντελοποίηση προσώπου με Ενεργά Μοντέλα Όψης (EMO)
 - Σχήμα και υφή αναλύονται σε χαμηλοδιάστατους υποχώρους
- **Οπτικά χαρακτηριστικά:** Παράμετροι του EMO $\mathbf{q} = [\mathbf{p}^T, \boldsymbol{\lambda}^T]^T$
- Προσαρμογή μοντέλου στον ομιλητή για εξαγωγή χαρακτηριστικών ομιλίας



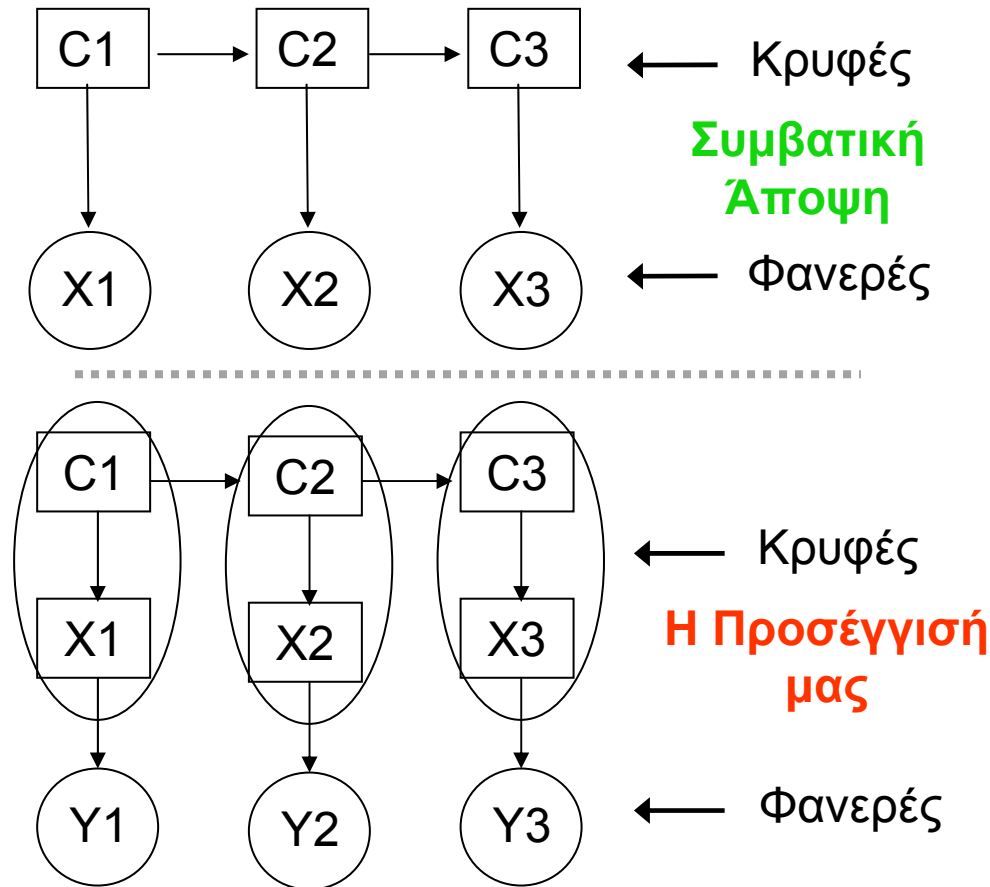
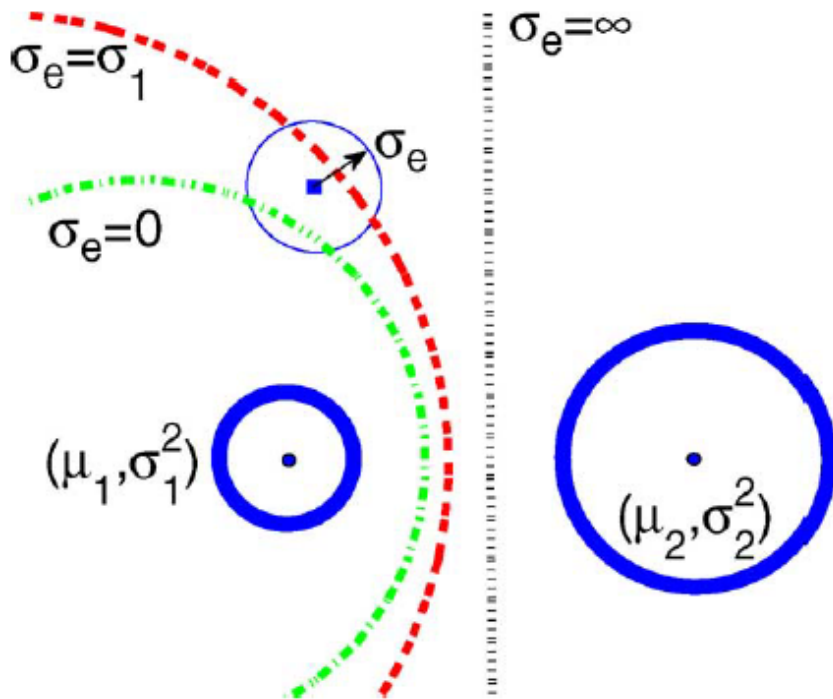
Σύστημα πραγματικού χρόνου για οπτική-ακουστική αναγνώριση ομιλίας

Demo σε IEEE συνέδρια: MMSP, ICASSP, CVPR
Ευχ.: Μ. Περακάκης, Α. Ποταμιάνος, Πολ. Κρήτης



Πολυτροπική σύμμιξη με αντιστάθμιση αβεβαιότητας

- Λαμβάνοντας υπόψη την αβεβαιότητα μέτρησης χαρακτηριστικών οδηγεί σε προσαρμοστικά πιθανοτικά μοντέλα σύμμιξης



Βασικές συνεισφορές διατριβής

- Ψηφιακή αποτύπωση τοιχογραφιών Ακρωτηρίου Θήρας με τεχνικές υπολογιστικής φωτογραφίας
- Ανάλυση εικόνας και όραση υπολογιστών
 - Πολυπλεγματική αριθμητική επίλυση μερικών διαφορικών εξισώσεων όρασης υπολογιστών. Εφαρμογές:
 - Ενδοσυμπλήρωση με μοντέλα ανισοτροπικής διάχυσης
 - Κατάτμηση με μοντέλα γεωμετρικών ενεργών περιγραμμάτων
 - Στατιστικά μοντέλα για ενδοσυμπλήρωση στο πεδίο των κυματιδίων
 - Βασική εφαρμογή: Αποκατάσταση αρχαίων τοιχογραφιών
- Συνδυασμένη οπτική-ακουστική μοντελοποίηση φωνής
 - Σύμμιξη πολυτροπικών σημάτων με αντιστάθμιση αβεβαιότητας
 - Μοντέλα όρασης υπολογιστών για περιγραφή προσώπου

Δημοσιεύσεις διατριβής

1. **G. Papandreou** and P. Maragos, *Multigrid Geometric Active Contour Models*, IEEE Transactions on Image Processing, vol. 16, no. 1, pp. 229-240, Jan. 2007
 2. **G. Papandreou**, A. Katsamanis, V. Pitsikalis, and P. Maragos, *Adaptive Multimodal Fusion by Uncertainty Compensation with Application to Audio-Visual Speech Recognition*, IEEE Transactions on Audio, Speech and Language Processing, vol. 17, no. 3, pp. 423-435, Mar. 2009
 3. A. Katsamanis, **G. Papandreou**, and P. Maragos, *Face Active Appearance Modeling and Speech Acoustic Information to Recover Articulation*, IEEE Transactions on Audio, Speech and Language Processing, vol. 17, no. 3, pp. 411-422, Mar. 2009
 4. S. Lefkimmiatis, P. Maragos, and **G. Papandreou**, *Bayesian Inference on Multiscale Models for Poisson Intensity Estimation: Applications to Photon-Limited Image Denoising*, IEEE Transactions on Image Processing, vol. 18, no. 8, pp. 1724-1741, Aug. 2009
 5. **G. Papandreou** and P. Maragos, *Image Inpainting with Transform Domain Hidden Markov Tree Models*, προετοιμάζεται για υποβολή
- + 13 δημοσιεύσεις σε διεθνή συνέδρια με κριτή
 - + 2 κεφάλαια σε βιβλίο

Ευχαριστώ πολύ!