

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ  
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ  
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

**Δυναμική Ανίχνευση και Εντοπισμός Κινούμενων  
Αντικειμένων σε Ακολουθίες Εικόνων**

Νικόλαος Κ. Παραγυός

Μεταπτυχιακή Εργασία

Ηράκλειο, 1996



ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ  
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ  
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

## Δυναμική Ανίχνευση και Εντοπισμός Κινοϋμένων Αντικειμένων σε Ακολουθίες Εικόνων

Εργασία που υποβλήθηκε από τον  
Νικόλαο Κ. Παραγυιό  
ως μερική εκπλήρωση των απαιτήσεων  
για την απόκτηση  
ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΕΙΔΙΚΕΥΣΗΣ

Συγγραφέας:

---

Νικόλαος Κ. Παραγυιός  
Τμήμα Επιστήμης Υπολογιστών

Εισηγητική Επιτροπή:

---

Γεώργιος Τζιρίτας  
Αναπληρωτής Καθηγητής, Επόπτης

---

Στέλιος Ορφανουδάκης  
Καθηγητής, Μέλος

---

Γεώργιος Γεωργακόπουλος  
Επίκουρος Καθηγητής, Μέλος

Δεκτή:

---

Πάνος Κωνσταντόπουλος  
Πρόεδρος Επιτροπής Μεταπτυχιακών Σπουδών

Ηράκλειο, 1996



*Στην Ειρήνη*



# Δυναμική Ανίχνευση και Εντοπισμός Κινούμενων Αντικειμένων σε Ακολουθίες Εικόνων

Νίκος Παραγυιός  
Μεταπτυχιακή Εργασία

Τμήμα Επιστήμης Υπολογιστών  
Πανεπιστήμιο Κρήτης

## Περίληψη

Το πρόβλημα της ανίχνευσης κινούμενων αντικειμένων σε ακολουθίες εικόνων αποτελεί βασική προϋπόθεση σε αρκετές εφαρμογές της Ανάλυσης Κίνησης, όπως η κωδικοποίηση. Στην περίπτωση στατικού παρατηρητή (κάμερα) είναι προφανές ότι οι όποιες αλλαγές παρατηρούνται οφείλονται σε κινούμενα αντικείμενα, ενώ διαφορετικά είναι δύσκολος ο διαχωρισμός των μεταβολών που οφείλονται στην κίνηση του παρατηρητή και στα ανεξάρτητα κινούμενα αντικείμενα. Το πρόβλημα αυτό μπορεί να αντιμετωπιστεί με τον υπολογισμό της κίνησης του παρατηρητή και την αντιστάθμιση της στην συνέχεια, στην ακολουθία εισόδου. Ένα επιπλέον πρόβλημα είναι ο προσδιορισμός της θέσης των αντικειμένων αυτών, αφού μια περιοχή μεταβολών μπορεί να οφείλεται: (1) *Επικάλυψη υποστρώματος (background)*, (2) *Αποκάλυψη υποστρώματος* (3) *Προβολές διαφορετικών τμημάτων των κινούμενων αντικειμένων*. Γνωρίζοντας τις παραπάνω περιοχές είναι δυνατός ο προσδιορισμός των αντικειμένων στο ζεύγος των εικόνων, αφού η ένωση των περιοχών (1) και (3) μας δίνει τις θέσεις τους στην πρώτη εικόνα, ενώ η ένωση των περιοχών (2) και (3) στην δεύτερη εικόνα.

Η εργασία αυτή αντιμετωπίζει τα παραπάνω προβλήματα, όπου η μοντελοποίηση τους γίνεται με την χρήση πεδίων Markov. Επιπλέον κατανομές Gibbs χρησιμοποιούνται για να περιγράψουν τους χάρτες αποφάσεων (detection maps). Η ανίχνευση κινούμενων αντικειμένων ανάγεται σε πρόβλημα εξομάλυνσης, όπου η παρατηρούμενη συνάρτηση πυκνότητας πιθανότητας της διαφοράς φωτεινότητας, θεωρείται μίξη δύο κατανομών Laplace, που εκφράζουν την υπόθεση κίνησης και την απουσία της. Επιπλέον προτείνεται μία νέα μέθοδος αυτόματου καθορισμού της παραμέτρου του προβλήματος εξομάλυνσης, βασισμένη σε τεχνική ψηφοδοσίας. Για το δεύτερο πρόβλημα, δημιουργείται ένα δυναμικό μοντέλο βασισμένο στις φωτεινότητες του ζεύγους των εικόνων όπου είτε με χρήση κατανομών Gauss είτε με χρήση τμηματικά ομοιόμορφων

κατανομών περιγράφονται στατιστικά οι τέσσερις υποθέσεις. Η λύση και των δύο προβλημάτων ανάγεται σε ελαχιστοποίηση συναρτήσεων κόστους, με την χρήση αιτιοκρατικών αλγορίθμων καλάρωσης, που είναι οι Επαναλαμβανόμενες Τοπικά Επικρατούσες Τιμές (ΕΤΕΤ) και η Μέγιστη Εμπιστοσύνη κατά Προτεραιότητα (ΜΕκΠ). Επιπλέον δύο διαφορετικές τεχνικές πολλαπλών επιπέδων εφαρμόζονται για μείωση του απαιτούμενου υπολογιστικού κόστους. Πειραματικά αποτελέσματα εξάγονται σε πραγματικές αλλά και συνθετικές ακολουθίες εικόνων, με στόχο τον έλεγχο της απόδοσης και της αξιοπιστίας των προτεινόμενων μεθόδων.

Από την εργασία αυτή προέκυψαν οι δημοσιεύσεις: [48], [49].

Επόπτης: Γεώργιος Τζιφίτας,  
Αναπληρωτής καθηγητής,  
Πανεπιστήμιο Κρήτης.



# Adaptive detection and location of moving objects in image sequences

Nikos Paragios

Master Thesis

Computer Science Department

University of Crete, Greece

## Abstract

Two important problems in motion analysis are addressed in this thesis. The detection of moving objects and their location. Indeed, complete motion detection is not equivalent to temporal change detection. Presence of motion usually causes three kinds of “change regions” to appear. They correspond to: (1) *Covered background*, (2) *Uncoverd static background*, (3) *An overlap of two successive object projection*. Moving objects can be located precisely by the use of these regions, where the union of (1) and (3) gives the object locations in second frame, while (2) and (3) in first frame.

A statistical approach is adopted to formulate each problem. For the first, the inter-frame difference is modeled as a mixture of Laplacian distributions, while a Gibbs random field is used for describing the label field. A new method to determine the regularization parameter is proposed, based on a voting technique. The solution of the second problem is based on the observation of two successive frames alone. Using the results of moving objects detection an adaptive statistical model for the couple of image intensities is identified. For both problems two different multi-scale algorithms are evaluated, while the labeling problem is solved using either ICM (Iterated Conditional Modes) or HCF (Highest Confidence First) algorithms. Experimental results on synthetic and real video sequences illustrate the efficiency of the proposed method.

**Advisor:** George Tziritas,  
Associate professor,  
University of Crete.



# Ευχαριστίες

Η παρούσα εργασία δεν θα μπορούσε να ολοκληρωθεί χωρίς την ουσιαστική συμβολή του επιβλέποντα καθηγητή μου. Ετσι ευχαριστώ θερμά τον κ. Γιώργο Τζιρίτα, ο οποίος μου έδωσε την δυνατότητα να ασχοληθώ με το θέμα αυτό και με καθοδήγησε σε όλη την διάρκεια της εργασίας μου. Ο σχεδιασμός και η υλοποίηση της εργασίας αυτής βασίστηκε σε δουλειά που είχε ξεκινήσει ο κ. Γιώργος Τζιρίτας τόσο σε θεωρητικό όσο και σε επίπεδο υλοποίησης. Σε όλη την διάρκεια της συνεργασίας μας ήταν πάντα ανοιχτός και πρόθυμος να με βοηθήσει στα διάφορα προβλήματα που προέκυπταν. Επιπλέον θέλω να τον ευχαριστήσω για την δυνατότητα που μου έδωσε να περάσω τέσσερις μήνες στο εξωτερικό, μέσα σε μια πολύ καλή ερευνητική ομάδα των ενδιαφερόντων μου. Ετσι χωρίς δόση υπερβολής και όχι γιατί συνηθίζεται για λόγους τυπικούς μετά το τέλος κάθε συνεργασίας, μπορώ να πω χωρίς αμφιβολία ότι φεύγω από το Ηράκλειο ευχόμενος να έχω την τύχη να συνεργαστώ με ανθρώπους σαν τον κ. Γιώργο Τζιρίτα και στο μέλλον.

Θα ήταν παράλειψη εκ μέρους μου να μην ευχαριστήσω τους γονείς μου, Ειρήνη και Κωνσταντίνο Παραγυιό που όλα αυτά τα χρόνια με στήριξαν υπερβάλλοντας εαυτούς, και θυσιάζοντας πολλές από τις δυνάμεις τους για μένα. Ετσι χωρίς αμφιβολία ήταν το σίγουρο λιμάνι στις δύσκολες στιγμές μου. Μου είναι πολύ δύσκολο να εκφράσω τα συναισθήματα μου μέσα σε λίγες μόνο λέξεις, αλλά θέλω να είναι σίγουροι ότι όλες αυτές οι θυσίες με έχουν σημαδέψει, και θα τους αγαπώ και θα τους εκτιμώ για πάντα. Επίσης θέλω να ευχαριστήσω τα αδέρφια μου Πέτρο και Βασίλη, για την αμέριστη υποστήριξη που μου έδωσαν καθ' όλη την διάρκεια των σπουδών μου.



# Περιεχόμενα

<b>Περίληψη</b>	<b>i</b>
<b>Abstract</b>	<b>iii</b>
<b>Ευχαριστίες</b>	<b>v</b>
<b>Περιεχόμενα</b>	<b>viii</b>
<b>Κατάλογος Σχημάτων</b>	<b>x</b>
<b>1 Γενική Εισαγωγή</b>	<b>1</b>
<b>2 Στατιστικά Μοντέλα</b>	<b>5</b>
2.1 Πεδία Markov . . . . .	5
2.2 Τεχνικές Πολλαπλών Επιπέδων . . . . .	8
2.2.1 Πολλαπλά Επίπεδα στον χώρο Δεδομένων και Αποφάσεων . . . . .	10
2.2.2 Πολλαπλά Επίπεδα στον Χώρο Αποφάσεων . . . . .	11
2.3 Απιοκρατικοί Αλγόριθμοι Χαλάρωσης . . . . .	12
2.3.1 Επαναλαμβανόμενες Τοπικά Επικρατούσες Τιμές . . . . .	13
2.3.2 Μέγιστη Εμπιστοσύνη κατά Προτεραιότητα . . . . .	14
<b>3 Εκτίμηση Παραμέτρου Εξομάλυνσης</b>	<b>19</b>
3.1 Εισαγωγή-Ανασκόπηση . . . . .	19
3.2 Εκτίμηση της παραμέτρου εξομάλυνσης με Ψηφοδοσία . . . . .	21
<b>4 Ανίχνευση Κινούμενων Αντικειμένων</b>	<b>25</b>
4.1 Εισαγωγή-Ανασκόπηση . . . . .	25

4.2	Εκτίμηση Ιδίας Κίνησης . . . . .	27
4.3	Ανίχνευση Ανεξάρτητα Κινούμενων Αντικειμένων . . . . .	30
<b>5</b>	<b>Εντοπισμός Θέσης Κινούμενων Αντικειμένων</b>	<b>39</b>
5.1	Εισαγωγή-Ανασκόπηση . . . . .	39
5.2	Στατιστική Ανάλυση . . . . .	41
5.2.1	Πιθανοτική προσέγγιση με τμηματικά ομοιόμορφες κατανομές . . . . .	42
5.2.2	Πιθανοτική προσέγγιση με κατανομές Gauss . . . . .	44
5.3	Μοντελοποίηση Προβλήματος-Αποτελέσματα . . . . .	46
<b>6</b>	<b>Συμπεράσματα-Μελλοντική Εργασία</b>	<b>51</b>
6.1	Συμπεράσματα . . . . .	52
6.1.1	Σύγκριση με Άλλες Προσεγγίσεις . . . . .	52
6.1.2	Τεχνική Ψηφοδοσίας . . . . .	52
6.1.3	Αλγόριθμοι Ελαχιστοποίησης . . . . .	54
6.1.4	Τεχνικές Πολλαπλών Επιπέδων . . . . .	57
6.2	Μελλοντική Εργασία . . . . .	58
<b>A</b>	<b>Κβαντιστής Lloyd-Max</b>	<b>61</b>
<b>B</b>	<b>Εκτιμητής παραμέτρων σε προβλήματα Μίξης Κατανομών</b>	<b>63</b>
	<b>Βιβλιογραφία</b>	<b>65</b>
	<b>Ευρετήριο Ορων</b>	<b>71</b>

# Κατάλογος Σχημάτων

2.1	Γειτονιά Δεύτερης Τάξης 8-σημείων . . . . .	6
2.2	Συνάρτηση δυναμικού: Τα σημεία με το ίδιο χρώμα αντιστοιχούν σε ίδια απόφαση	7
2.3	Πολλαπλά επίπεδα στην ανακατασκευή: Οι αποφάσεις αναπαριστώνται με διανύσματα για ευκολότερη κατανόηση της μεθόδου . . . . .	8
2.4	Διαδικασία Παρεμβολής . . . . .	9
2.5	Πολλαπλά Επίπεδα στο Χώρο Δεδομένων & Αποφάσεων . . . . .	10
2.6	Πολλαπλά Επίπεδα στον Χώρο Αποφάσεων . . . . .	11
2.7	Προβολή των σχέσεων παρέας στον πλήρη χώρο δεδομένων . . . . .	12
2.8	Επαναλαμβανόμενες Τοπικά Επικρατούσες Τιμές . . . . .	14
2.9	Μέγιστης Εμπιστοσύνη κατά Προτεραιότητα . . . . .	17
3.1	Εκτίμηση Παραμέτρου Εξομάλυνσης με Ψηφοδοσία . . . . .	23
4.1	Εκτίμηση Ιδίας Κίνησης για την ακολουθία <i>Interview</i> Παράμετροι Ιδίας Κίνησης: $(\alpha_{1,1}, \alpha_{2,1}) = (5.75, -0.3)$ (α) Εικόνα διαφοράς φωτεινότητας. (β) Εικόνα διαφοράς φωτεινότητας μετά την αφαίρεση της Ιδίας Κίνησης. . . . .	29
4.2	Ανάλυση Ιστογράμματος Διαφοράς Φωτεινότητας με κατανομές Laplace Ακολουθίες (α) <i>Trevor White</i> , (β) <i>Interview</i> (μετά την αφαίρεση της ίδιας κίνησης) . . . . .	31
4.3	Ανίχνευση Κινούμενων Αντικειμένων: Ακολουθία <i>Shpera</i> (Συνθετική) Πάνω : Εικόνες Εισόδου, Αρχική Εκκίνηση Κάτω : ΕΤΕΠ & ΠΕ (Δεδομένα & Αποφάσεις), ΜΕκΠ & ΠΕ (Αποφάσεις), Βέλτιστο Αποτέλεσμα . . . . .	33
4.4	Ανίχνευση Κινούμενων Αντικειμένων: Ακολουθία <i>Trevor White</i> Πάνω : Εικόνες Εισόδου, Αρχική Εκκίνηση Κάτω : ΕΤΕΠ & ΠΕ (Δεδομένα & Αποφάσεις), ΜΕκΠ & ΠΕ (Αποφάσεις), Αποτέλεσμα . . . . .	34
4.5	Διάγραμμα ψηφοδοσίας για την επιλογή του $\lambda$ . . . . .	35

4.6	Δυναμική Ανίχνευση Κινούμενων Αντικειμένων (α) Ακολουθία <i>Highway</i> (ΕΤΕΤ), $\lambda=1.0125$ (β) Ακολουθία <i>Interview</i> - Κινούμενος Παρατηρητής (ΜΕκΠ), $\lambda=0.8125$ (γ) Ακολουθία <i>Kollnig</i> (ΜΕκΠ), $\lambda=0.8625$ (δ) Ακολουθία <i>Van</i> (ΕΤΕΤ), $\lambda=0.9875$ . . . . .	37
4.7	Ανίχνευση Κινούμενων Αντικειμένων, $\lambda=1.00$ (α) Ακολουθία <i>Highway</i> (ΕΤΕΤ) (β) Ακολουθία <i>Interview</i> - Κινούμενος Παρατηρητής (ΜΕκΠ) (γ) Ακολουθία <i>Kollnig</i> (ΜΕκΠ) (δ) Ακολουθία <i>Van</i> (ΕΤΕΤ) . . . . .	38
5.1	Ακολουθία <i>Trevor White</i> - Ανάλυση Ιστογράμματος Τμηματικά Ομοιόμορφες Κατανομές: (α) Στατικό Τμήμα, (β) Κινούμενο Τμήμα . . . . .	43
5.2	Ακολουθία <i>Trevor White</i> - Ανάλυση Ιστογράμματος Κατανομές Gauss : (α) Στατικό Τμήμα, (β) Κινούμενο Τμήμα . . . . .	44
5.3	Ανίχνευση περιοχών κίνησης στην ακολουθία <i>Sphera</i> (α) Πιθανοτική προσέγγιση με τμηματικά ομοιόμορφες κατανομές (β) Πιθανοτική προσέγγιση με χρήση κατανομών Gauss . . . . .	47
5.4	Εντοπισμός Θέσης Κινούμενων Αντικειμένων για την ακολουθία <i>Trevor White</i> (α) Ανάλυση ιστογράμματος με κατανομές Gauss (β) Ανάλυση ιστογράμματος με τμηματικά ομοιόμορφες κατανομές (1) Τμηματοποίηση με κριτήριο Μέγιστης Πιθανοφάνειας (2) Τμηματοποίηση με χρήση Πολλαπλών επιπέδων και ΕΤΕΤ (3) Τμηματοποίηση με χρήση Πολλαπλών επιπέδων και ΜΕκΠ . . . . .	49
5.5	Εντοπισμός θέσης Κινούμενων Αντικειμένων στο ζεύγος των εικόνων . . . . .	50
6.1	Συγκριτικά Αποτελέσματα, Ακολουθία (α) Εικόνες Ακολουθίας, (β) Ανίχνευση Ανεξάρτητα Κινούμενων Αντικειμένων: Με χρήση πεδίων Βάθους και 3-Δ κίνησης (αριστερά), Προτεινόμενη προσέγγιση (δεξιά) . . . . .	53
6.2	Αυτόματος καθορισμός της παραμέτρου εξομάλυνσης με μη βέλτιστη ανακατασκευή: (α) Ακολουθία <i>Trevor White</i> , (β) Ακολουθία <i>Sphera</i> . . . . .	54
6.3	Υπολογιστικό κόστος: (α) Ένα Επίπεδο, (β) Πολλαπλά Επίπεδα . . . . .	55
6.4	Ευστοχησία στην Παράμετρο Εξομάλυνσης: (α) Ένα Επίπεδο, (β) Πολλαπλά Επίπεδα . . . . .	56
6.5	Εξάρτηση από την Λύση Μέγιστης Πιθανοφάνειας: (α) Ένα Επίπεδο, (β) Πολλαπλά Επίπεδα . . . . .	57



# Κεφάλαιο 1

## Γενική Εισαγωγή

Το πρόβλημα της ανίχνευσης ανεξάρτητα κινούμενων αντικειμένων καθώς και του προσδιορισμού της θέσης τους, αποτελεί αντικείμενο μελέτης για πολλές εφαρμογές στον τομέα της ανάλυσης εικόνων ([15], [28] και [39]). Επίσης το ίδιο πρόβλημα εμφανίζεται στην κωδικοποίηση και μετάδοση ακολουθιών εικόνων ([21], [43] και [57]). Η λύση του είναι ευκολότερη στην περίπτωση που τα δεδομένα προέρχονται από στατικό παρατηρητή, και αρκετά διαφορετικές προσεγγίσεις έχουν προταθεί κάτω από αυτή την συνθήκη ([2], [28] και [42]), βασισμένες στη διαφορά των φωτεινιοτήτων ανά σημείο στο ζεύγος των εικόνων. Οι προσεγγίσεις αυτές ποικίλουν και συναντάμε από πολύ απλές τεχνικές κατωφλιού μέχρι πολύπλοκες μεθόδους, όπου η λύση του προβλήματος προέρχεται από την ελαχιστοποίηση πολύπλοκων συναρτήσεων κόστους. Ορισμένες από τις προσεγγίσεις αυτές μπορούν να αντιμετωπίσουν και μικρές μετακινήσεις (ταλαντώσεις) του παρατηρητή ([42]).

Στην περίπτωση κινούμενου παρατηρητή η λύση του προβλήματος είναι αρκετά πιο δύσκολη, αφού δεν είναι εύκολος ο διαχωρισμός των αλλαγών που οφείλονται στην κίνηση του παρατηρητή και αυτών που οφείλονται στην κίνηση των ανεξάρτητα κινούμενων αντικειμένων. Για να αντιμετωπιστεί το πρόβλημα αυτό απαιτείται αρχικά η εκτίμηση της ίδιας κίνησης του παρατηρητή, με την χρήση της οποίας είναι εφικτή η δημιουργία μιας αντισταθμισμένης ακολουθίας εικόνων, όπου το υπόστρωμα παραμένει αμετάβλητο (ακίνητο), οπότε οι όποιες μεταβολές παρατηρούνται οφείλονται στην κίνηση ανεξάρτητων αντικειμένων.

Η εργασία αυτή αντιμετωπίζει τα παραπάνω προβλήματα, την ανίχνευση κινούμενων αντικειμένων καθώς και τον προσδιορισμό της θέσης τους. Είναι γεγονός ότι η ανίχνευση μεταβολών σε ένα ζεύγος εικόνων δεν είναι ισοδύναμη με την πλήρη ανίχνευση των κινούμενων αντικειμένων. Η παρουσία ανεξάρτητης κίνησης συνήθως δημιουργεί τρεις διαφορετικές περιοχές στην εικόνα σχετιζόμενες με αυτή. Ετσι υπάρχουν περιοχές που αντιστοιχούν:

- (1) *Επικάλυψη υποστρώματος από ανεξάρτητα κινούμενα αντικείμενα,*
- (2) *Αποκάλυψη*

υποστρώματος εξαιτίας της κίνησης αυτών των αντικειμένων, και (3) *Επικάλυψη τμημάτων του κινούμενου αντικειμένου*.

Θα πρέπει να προσθέσουμε ότι οι περιοχές που αντιστοιχούν στην τρίτη κατηγορία είναι πολύ δύσκολο να εντοπιστούν στις περιπτώσεις που οι φωτεινότητες στα αντικείμενα είναι ομοιόμορφες, οπότε η κίνηση τους δεν επιφέρει σημαντικές τιμές στην διαφορά τους από εικόνα σε εικόνα. Αντίθετα ο εντοπισμός των δύο άλλων κατηγοριών είναι συνήθως πιο εύκολος. Αν στις παραπάνω υποθέσεις προσθέσουμε την απουσία κίνησης, (4) *Επικάλυψη του στατικού τμήματος των εικόνων*, δημιουργείται ένα πλήρες μοντέλο ποιοτικού προσδιορισμού της κίνησης, το οποίο αποτελεί επέκταση του μοντέλου ανίχνευσης μεταβολών, και επιπλέον υπολογισμοί απαιτούνται για την περιγραφή του. Η ύπαρξη αυτού του είδους της πληροφορίας μας δίνει την δυνατότητα προσδιορισμού της θέσης των αντικειμένων στο ζεύγος των εικόνων. Έτσι η ένωση των περιοχών (1) και (3) μάς δίνει την θέση των αντικειμένων στη δεύτερη εικόνα, ενώ από την ένωση των περιοχών (2) και (3) προκύπτει η θέση των αντικειμένων στη πρώτη εικόνα.

Στην εργασία αυτή, αντιμετωπίζονται και τα δύο προβλήματα με χρήση στατιστικών μοντέλων. Πεδία Markov χρησιμοποιούνται για το λόγο αυτό και η περιγραφή του χάρτη αποφάσεων γίνεται με κατανομές Gibbs. Συναρτήσεις κόστους κατασκευάζονται βασιζόμενες σε αυτές τις κατανομές όπου μη γραμμικές αλληλεπιδράσεις ορίζονται μεταξύ των διαθέσιμων παρατηρήσεων και της ανακατασκευαζόμενης τμηματοποίησης. Οι συναρτήσεις αυτές έχουν μορφή που ανάγει τα προβλήματα σε τεχνικές εξομάλυνσης. Μία νέα μέθοδος για την αυτόματη επιλογή της παραμέτρου εξομάλυνσης προτείνεται, βασιζόμενη σε τεχνική ψηφοδοσίας. Το πρόβλημα της ανακατασκευής διαχωρίζεται με την εκτίμηση της βέλτιστης τιμής για την παράμετρο εξομάλυνσης. Η ιδέα είναι η χρήση ενός εναλλακτικού σχήματος όπου η λύση του ενός προβλήματος οδηγεί σε μία πιο σταθερή εκτίμηση της λύσης του άλλου. Η διαδικασία αυτή επαναλαμβάνεται διαδοχικά μέχρις ότου να προκύψει μία ικανοποιητική εκτίμηση του χάρτη αποφάσεων.

Για την ανίχνευση κινούμενων αντικειμένων, η παρατηρούμενη κατανομή πιθανοτήτων των μεταβολών στην φωτεινότητα προσεγγίζεται σαν μίξη δύο κατανομών Laplace (μία που εκφράζει την υπόθεση κίνησης, και μία την απουσία της). Οι παράμετροι των κατανομών αυτών εκτιμώνται με την χρήση κριτηρίου μέγιστης πιθανοφάνειας. Για τον προσδιορισμό της θέσης τους, είτε τμηματικά ομοιόμορφες κατανομές, είτε διοδιάστατες κατανομές Gauss χρησιμοποιούνται για να περιγράψουν στατιστικά το ζεύγος των φωτεινοτήτων ανά σημείο. Οι χάρτες αποφάσεων προκύπτουν από την ελαχιστοποίηση των οριζόμενων συναρτήσεων κόστους, με κριτήριο την μεγιστοποίηση της εκ των υστέρων πιθανότητας (ΜΥΠ). Το μεγάλο υπολογιστικό κόστος που απαιτείται οδηγεί σε χρήση τεχνικών πολλαπλών επιπέδων, με την οποία επιτυγχάνεται σημαντική μείωση του, ενώ δύο διαφορετικοί αυτοκρατικοί αλγόριθμοι καλάρωσης χρησιμοποιούνται για την ελαχιστοποίηση στα διάφορα επίπεδα. Οι αλγόριθμοι

αυτοί παρουσιάζουν μικρό υπολογιστικό κόστος, σε σχέση με τους στοχαστικούς και είναι ο ΕΤΕΤ (Επαναλαμβανόμενες Τοπικά Επικρατούσες Τιμές) και ο ΜΕκΠ (Μέγιστη Εμπιστοσύνη κατά Προτεραιότητα), με την χρήση των οποίων τελικά προκύπτουν οι χάρτες αποφάσεων.

Ο περιορισμός για στατικό παρατηρητή (κάμερα) αντιμετωπίζεται επιτυχώς με την εκτίμηση της ίδιας (κυρίαρχης) κίνησης, με αλγόριθμο σε πολλαπλά επίπεδα. Η χρήση πολλαπλών επιπέδων μας δίνει την δυνατότητα αντιμετώπισης περιπτώσεων όπου υπάρχουν μεγάλες μετακινήσεις, που είναι εξαιρετικά δύσκολο να εκτιμηθούν με την χρήση αλγορίθμων ενός επιπέδου. Η ίδια κίνηση περιγράφεται με την χρήση γραμμικών μοντέλων (αφφινικό μοντέλο 6-παραμέτρων) και στη συνέχεια πραγματοποιείται αντιστάθμιση της, με αποτέλεσμα η θέση του υποστρώματος παραμένει αμετάβλητη.

Η εργασία αυτή ακολουθεί την εξής δομή. Στο Κεφάλαιο 2 υπάρχει συνοπτική περιγραφή των στατιστικών μοντέλων που χρησιμοποιούνται καθώς και των αντίστοιχων αλγορίθμων, ενώ στο Κεφάλαιο 3 εξετάζεται το πρόβλημα της βέλτιστης εκτίμησης της παραμέτρου εξομάλυνσης με μικρή ανασκόπηση των ήδη υπαρχόντων προσεγγίσεων. Το Κεφάλαιο 4 αντιμετωπίζει το πρόβλημα της ανίχνευσης κινούμενων αντικειμένων. Αρχικά προτείνεται το μοντέλο για την εκτίμηση της ίδιας κίνησης. Η θέση των αντικειμένων στο ζεύγος των εικόνων αποτελεί το αντικείμενο μελέτης του Κεφαλαίου 5. Τα συμπεράσματα καθώς και οι πιθανές επεκτάσεις τις εργασίας αυτής παρατίθενται στο Κεφάλαιο 6, όπου υπάρχουν και συγκριτικά αποτελέσματα όσον αφορά τις διαφορετικές μεθόδους που χρησιμοποιήθηκαν. Στο Παράρτημα Α παρατίθενται οι συνθήκες που διέπουν την κατασκευή κβαντιστή κατά τους Lloyd-Max, ενώ στο Παράρτημα Β αναλύονται οι εξισώσεις εκτίμησης παραμέτρων μίξης ιδίου τύπου κατανομών, βαισιζόμενες σε κριτήριο μέγιστης πιθανοφάνειας. Πειραματικά αποτελέσματα παρατίθενται σε όλα τα τμήματα της εργασίας.

Με στόχο τον έλεγχο της απόδοσης και της αξιοπιστίας των προτεινόμενων μεθόδων, πειραματικά αποτελέσματα εξάγονται με την χρήση συνθετικών και πραγματικών δεδομένων. Στα δεδομένα εισόδου υπάρχουν ακολουθίες χωρίς ίδια κίνηση αλλά και ακολουθίες που προέρχονται από κινούμενες κάμερες με ταυτόχρονη παρουσία ανεξάρτητα κινούμενων αντικειμένων.



## Κεφάλαιο 2

# Στατιστικά Μοντέλα

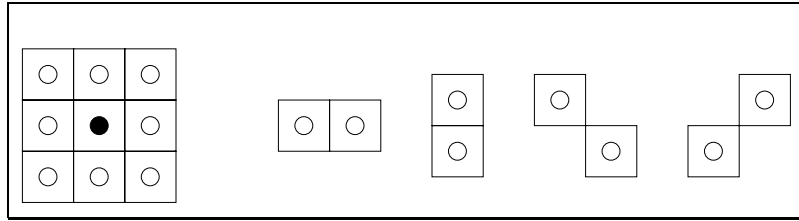
### 2.1 Πεδία Markov

Τα πεδία Markov έχουν χρησιμοποιηθεί σε πολλές εφαρμογές για μοντελοποίηση προβλημάτων ανακατασκευής τα τελευταία χρόνια ([10], [22], [24], [30] και [34]). Για την καλύτερη κατανόηση τους, απαιτούνται κάποιοι ορισμοί που σχετίζονται με τον συμβολισμό που θα χρησιμοποιηθεί στη συνέχεια. Εστω  $S = \{s_1, s_2, \dots, s_N\}$  το σύνολο των σημείων της εικόνας και  $L = \{l_1, l_2, \dots, l_Q\}$  σύνολο αμοιβαία αποκλεισμένων αποφάσεων που σχετίζονται με την ανακατασκευή. Εστω  $X_s$  τυχαία μεταβλητή για το  $s \in S$ , με τιμές στο σύνολο  $L$ . Ο χάρτης αποφάσεων  $\omega$  είναι ένα σύνολο τυχαίων μεταβλητών  $X = \{X_s, s \in S\}$ . Εστω  $\omega_s = \omega(X_s) \in L$  η απόφαση που αντιστοιχεί στο σημείο  $s$  σύμφωνα με τον χάρτη αποφάσεων  $\omega$ , και  $\Omega$  το σύνολο όλων των πιθανών ανακατασκευών σε σχέση με τα  $S$  και  $L$ . Με βάση τους παραπάνω ορισμούς για το πρόβλημα της ανακατασκευής έχουμε:

- Ένα σύνολο σημείων (ή αντικειμένων)  $S$
- Ένα σύνολο πιθανών αποφάσεων για κάθε σημείο  $s \in S$ ,  $X_s$  με τιμές στο  $L$
- Επιπλέον μπορούμε να ορίσουμε ένα σύστημα γειννίας  $\Gamma$ , το οποίο ορίζει γράφο όπου οι κόμβοι συμβολίζουν σημεία (αντικείμενα), και οι ακμές (με βάρη) τους περιορισμούς που υπάρχουν για τις αποφάσεις που λαμβάνονται για γειτονικά σημεία

Το πρόβλημα είναι η λήψη μίας απόφασης σε κάθε σημείο, ώστε αυτή να είναι συμβατή με τους περιορισμούς ([62]). Εστω  $E$  σύνολο μη διατεταγμένων ζευγών  $(s_i, s_j)$ , αντιστοιχώντας σε συσχετίσεις μεταξύ των σημείων του  $S$ . Το σύστημα γειννίας  $\Gamma = \{N_s | s \in S\}$ , όπου  $N_s$  είναι η γειτονιά του  $s$ , με την παρακάτω έννοια

1.  $s \notin N_s$



Σχήμα 2.1: Γειτονιά Δεύτερης Τάξης 8-σημείων

2.  $r \in N_s$ , αν, και μόνο αν,  $(s, r) \in E$

Με βάση τα παραπάνω,  $X$  είναι τυχαίο πεδίο Markov σε σχέση με το σύστημα γειννιάσης  $\Gamma$  και συνάρτηση πυκνότητας πιθανότητας  $P$ , αν, και μόνο αν:

- θετικότητα (*Positivity*)

$$P(X = \omega) > 0 \text{ για όλα τα } \omega \in \Omega \quad (2.1)$$

- Markovιότητα (*Markovianity*)

$$P(X_s = \omega_s | X_r = \omega_r, r \in S, r \neq s) = P(X_s = \omega_s | X_r = \omega_r, r \in N_s) \quad (2.2)$$

Η χρησιμότητα των πεδίων Markov σε προβλήματα ανακατασκευής, είναι ότι η εκ των προτέρων γνώση που έχουμε για εξαρτήσεις μεταξύ διαφορετικών οντοτήτων της ανακατασκευής και των διαθέσιμων παρατηρήσεων, μπορούν να μοντελοποιηθούν με την τεχνική των περιορισμών στις αποφάσεις για γειτονικά σημεία. Συνήθως οι οντότητες στις εικόνες είναι ομαλά δομημένες και οι εκ των προτέρων κατανομές που τις περιγράφουν είναι ομογενείς και ισοτροπικές. Το ακόλουθο θεώρημα, συμπληρώνει τον ορισμό των τυχαίων πεδίων Markov.

Θεώρημα HAMMERLSEY-CLIFFORD: Ένα τυχαίο πεδίο  $X$  είναι πεδίο Markov σε σχέση με ένα σύστημα γειννιάσης  $\Gamma$ , αν, και μόνο αν, υπάρχει συνάρτηση  $V_c$  τέτοια ώστε

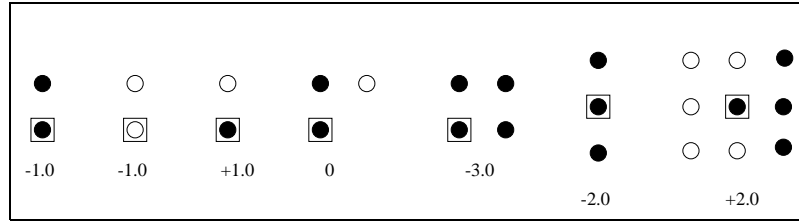
$$P(\omega) = \frac{1}{Z} e^{-\frac{1}{T} U(\omega)}, \forall \omega \in \Omega \quad (2.3)$$

όπου  $T$  και  $Z$  σταθερές και

$$U(\omega) = \sum_{c \in C} V_c(\omega) \quad (2.4)$$

όπου  $U(\omega)$  συνάρτηση κόστους (ενέργειας), και  $C$  το σύνολο των συνδεδεμένων σημείων (παρέες) σε σχέση με το σύστημα γειννιάσης  $\Gamma$  (Σχήμα 2.1). Τέλος  $V_c(\omega)$  συνάρτηση δυναμικού,

$$V_c(\omega_s, \omega_t) = \begin{cases} -\xi, & \text{αν } \omega_s = \omega_t \\ +\xi, & \text{αν } \omega_s \neq \omega_t \end{cases} \quad (2.5)$$



Σχήμα 2.2: Συνάρτηση δυναμικού: Τα σημεία με το ίδιο χρώμα αντιστοιχούν σε ίδια απόφαση

Στο Σχήμα 2.2 έχουμε ορισμένες τιμές αυτής της συνάρτησης για δοσμένη τμηματοποίηση στην γειτονιά κάποιου σημείου.

Η συνάρτηση πυκνότητας πιθανότητας που ορίζεται από τις σχέσεις (2.3) και (2.4) ονομάζεται κατανομή Gibbs. Η ισοδυναμία μεταξύ πεδίων Markov και κατανομών Gibbs δεν συσχετίζει απλώς τις τοπικά δεομευμένες συναρτήσεις πυκνότητας πιθανότητας με την ολική συνάρτηση πυκνότητας πιθανότητας, αλλά μας δίνει και τη δυνατότητα εύκολου προσδιορισμού τους, με την χρήση δυναμικών. Έτσι χρησιμοποιώντας το θεώρημα των HAMMERLSEY-CLIFFORD, και την εξίσωση (2.2) τα τοπικά χαρακτηριστικά μπορούν να υπολογιστούν μέσω της συνάρτησης δυναμικού, και πιο συγκεκριμένα ισχύει

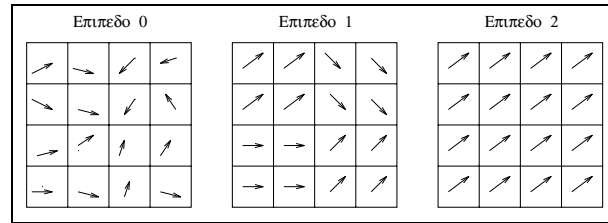
$$P(X_s = \omega_s | X_r = \omega_r, r \neq s) = \frac{e^{-\frac{1}{T} \sum_{c \in C_s} V_c(\omega)}}{\sum_{\omega'} e^{-\frac{1}{T} \sum_{c \in C_s} V_c(\omega')}} \quad (2.6)$$

όπου  $V_c$  συνάρτηση δυναμικού και  $C_s$  οι δυνατές παρέες στην γειτονιά του  $s$  με βάση το  $\Gamma$ .

Με τον ορισμό που έχει δοθεί μέχρι τώρα για τα πεδία Markov και την ισοδυναμία τους με τις κατανομές Gibbs, έχουν οριστεί αλληλεπιδράσεις μεταξύ γειτονικών σημείων, αλλά δεν υπάρχει η δυνατότητα να μοντελοποιήσουμε την εκ των προτέρων γνώση που έχουμε για συμβατότητα μεταξύ των αποφάσεων και των διαθέσιμων παρατηρήσεων. Κάτι τέτοιο έχει αποδειχθεί ότι μπορεί να γίνει ([34] και [20]) με αποτέλεσμα να προκύψει μία νέα συνάρτηση κόστους, της μορφής

$$U(\omega, d) = \sum_{c \in C_s} V_c(\omega) - T \sum_{s \in S} \ln[\delta(\omega_s, d_s)] \approx U_1(\omega) + U_2(\omega, d) \quad (2.7)$$

όπου η συνάρτηση  $\delta(\omega_s, d_s)$  δίνει ένα μέτρο της ποιότητας της ανακατασκευής σε σχέση με τις διαθέσιμες παρατηρήσεις. Με όμοιο τρόπο μπορούν να οριστούν και πιο πολύπλοκες μορφές για την συνάρτηση κόστους, με στόχο την βελτιστοποίηση του μοντέλου, αλλά και κύριο μειονέκτημα τη σημαντική αύξηση της πολυπλοκότητας του. Συνήθως η λύση του προβλήματος προκύπτει με κριτήριο τη μεγιστοποίηση της εκ των υστέρων πιθανότητας (ΜΥΠ), που είναι ισοδύναμη με την ελαχιστοποίηση της παραπάνω συνάρτησης κόστους  $U(\omega, d)$ .



Σχήμα 2.3: Πολλαπλά επίπεδα στην ανακατασκευή: Οι αποφάσεις αναπαριστώνται με διανύσματα για ευκολότερη κατανόηση της μεθόδου

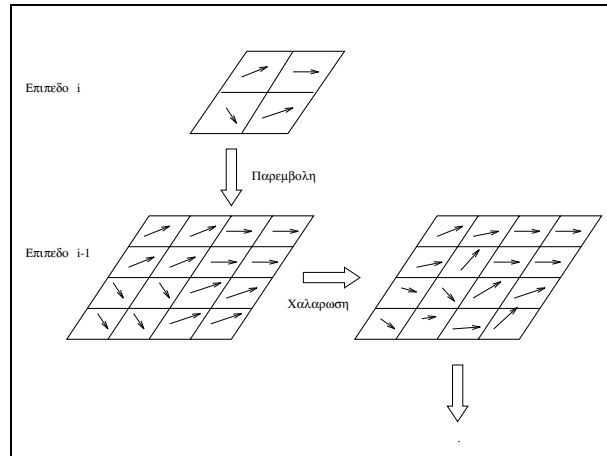
## 2.2 Τεχνικές Πολλαπλών Επιπέδων

Το πρόβλημα της ελαχιστοποίησης συναρτήσεων κόστους, όπως αυτές προκύπτουν από κατανομές Gibbs είναι συνήθως αρκετά πολύπλοκο, αφού ο αριθμός των δυνατών λύσεων είναι πολύ μεγάλος και επιπλέον οι συναρτήσεις αυτές μπορεί να παρουσιάζουν πολλά τοπικά ελάχιστα. Συνήθως στοχαστικοί αλγόριθμοι χαλάρωσης ([34]) χρησιμοποιούνται για την εξαγωγή της βέλτιστης λύσης, με σημαντικό μειονέκτημα το μεγάλο υπολογιστικό κόστος. Από την άλλη πλευρά έχει αποδειχθεί ότι οι τεχνικές πολλαπλών επιπέδων συμβάλλουν σημαντικά στην αύξηση της ταχύτητας σύγκλισης των διάφορων επαναληπτικών τεχνικών χαλάρωσης ([56]). Επίσης είναι γνωστό ότι οι τεχνικές αυτές είναι πολύ χρήσιμες στην περίπτωση που η συνάρτηση κόστους παρουσιάζει πολλά τοπικά ελάχιστα, κάτι που εμφανίζεται συχνά στην περίπτωση μη γραμμικών μοντέλων. Ετσι κυριαρχεί η άποψη ότι με την ανάλυση σε πολλαπλά επίπεδα συναρτήσεων κόστους, επιτυγχάνεται και η λείανση τους στα υψηλότερα επίπεδα, γεγονός που δίνει μερική λύση στο πρόβλημα της ύπαρξης τοπικών ελαχίστων ([3] και [51]).

Οι τεχνικές αυτές έχουν χρησιμοποιηθεί συχνά μέχρι τώρα για προβλήματα ανάλυσης περιεχομένου εικόνων, με κριτήριο είτε μερικές διαφορικές εξισώσεις ([32]), είτε πεδία Markov ([7]). Τα κυριότερα προβλήματα που εμφανίζονται κατά την χρήση τους, είναι η μορφή των συναρτήσεων κόστους στα διάφορα επίπεδα καθώς και οι τιμές των παραμέτρων τους, που συνήθως πρέπει να προσαρμόζονται ανάλογα με το επίπεδο επεξεργασίας. Μία συνηθισμένη προσέγγιση είναι η χρήση της ίδιας συνάρτησης σε όλα τα επίπεδα, αν και πολλές φορές οι αλληλεπιδράσεις μεταξύ των παρατηρήσεων και των αποφάσεων είναι άμεσα εξαρτώμενες από το επίπεδο επεξεργασίας. Ετσι σε πολλές προσεγγίσεις γίνεται αυτόματη αναπροσαρμογή της συνάρτησης κόστους στα διάφορα επίπεδα, γεγονός που μειώνει δραστικά την πολυπλοκότητα του μοντέλου.

Ετσι αντί να ορίζουμε την συνάρτησης κόστους απευθείας στον πλήρη χώρο λύσης του προβλήματος  $\Omega$ , που είναι και αρκετά μεγάλος, θεωρούμε διαδοχικά φωλιασμένα υποσύνολα  $\Omega_i$  που αποτελούνται από αποφάσεις που αντιστοιχούν σε μπλοκ σημείων μεγέθους  $2^i \times 2^i$ . Αυτοί οι φωλιασμένοι χώροι λύσεων ( $\Omega_i \subset \Omega_{i-1} \subset \dots \subset \Omega_1 \subset \Omega_0 = \Omega$ ) σχετίζονται με την περιγραφή της ανακατασκευής στα διάφορα επίπεδα. Το επίπεδο 0 αντιστοιχεί στον αυθεντικό



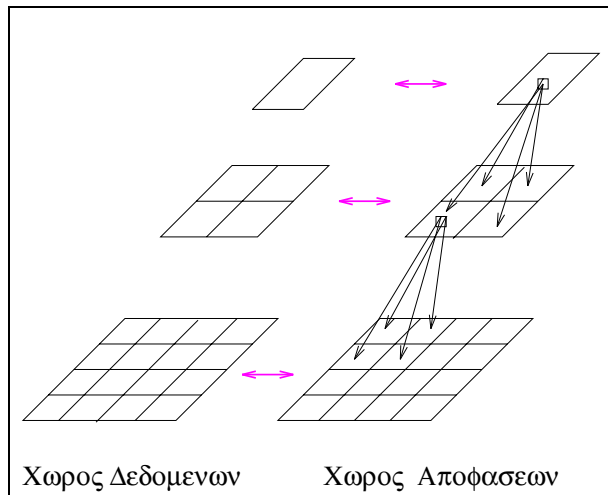


Σχήμα 2.4: Διαδικασία Παρεμβολής

χώρο λύσεων του προβλήματος  $\Omega$ , ενώ το επίπεδο  $i$  αντιστοιχεί σε λύσεις που είναι σταθερές πάνω σε κελιά (μπλοκ) διαστάσεων  $2^i \times 2^i$  (Σχήμα 2.3).

Είναι εύκολο να δει κανείς ότι στο επίπεδο  $i$  ο αριθμός των ανεξάρτητων αποφάσεων μειώνεται κατά ένα συντελεστή  $2^i \times 2^i$  οπότε η ανακατασκευή  $\Omega_i \subset \Omega = \{\omega_s, s \in S\}$  μπορεί να αναπαρασταθεί σε πλέγμα  $S^i$  μεγέθους  $2^{m-i} \times 2^{n-i}$  (αν υποθέσουμε ότι το αρχικό μέγεθος των εικόνων είναι  $2^m \times 2^n$ ). Το πρόβλημα στη συνέχεια ανάγεται στον τρόπο εκμετάλλευσης της γνώσης που υπάρχει σε κάποιο επίπεδο με την μετακίνηση στο επόμενο επίπεδο, όπου υπάρχει μεγαλύτερη ανάλυση. Ετσι υπάρχουν δύο διαφορετικές τεχνικές όσον αφορά τον τρόπο επεξεργασίας των πυραμίδων που δημιουργούνται. Η πιο συνηθισμένη είναι η μονόδρομη πυραμίδα, όπου αρχίζοντας από επίπεδα χαμηλής ανάλυσης και χρησιμοποιώντας διαδικασία προβολής των αποτελεσμάτων διατρέχουμε την πυραμίδα μέχρι το τελευταίο επίπεδο όπου έχουμε και τον αυθεντικό χώρο λύσης του προβλήματος (Σχήμα 2.4). Ετσι αφού εξασφαλιστεί η σύγκλιση του αλγορίθμου σε κάποιο επίπεδο, όπου μία απόφαση αντιστοιχεί σε παραπάνω από ένα σημεία στο επόμενο επίπεδο, τα αποτελέσματα προβάλλονται σε αυτό αντιγράφοντας απλώς την απόφαση σε όλα τα σημεία του αντίστοιχου μπλοκ. Στην συνέχεια δημιουργείται μια νέα συνάρτηση κόστους με βάση τις αρχικές αποφάσεις που έχουν προκύψει από το προηγούμενο επίπεδο, η οποία ελαχιστοποιείται (Σχήμα 2.4). Όταν εξασφαλιστεί η σύγκλιση και σε αυτό το επίπεδο προχωράμε στο επόμενο μέχρι να φτάσουμε στον αυθεντικό χώρο λύσης του προβλήματος ([36]).

Μία πιο ευέλικτη και πιο αξιόπιστη προσέγγιση, είναι η χρήση αμφίδρομης πυραμίδας όπου είναι δυνατή η κίνηση και προς τις δύο κατευθύνσεις. Ετσι με βάση κάποιο κριτήριο στην ποιότητα της σύγκλισης, είναι δυνατή η μετακίνηση στο προηγούμενο επίπεδο, αν η λύση που έχει προκύψει μετά την σύγκλιση δεν μας ικανοποιεί. Εκεί γίνεται αρχικά προβολή των τελευταίων αποτελεσμάτων (αντίθετη με αυτή που υπάρχει στην μονόδρομη πυραμίδα), και



Σχήμα 2.5: Πολλαπλά Επίπεδα στο Χώρο Δεδομένων & Αποφάσεων

αναζητείται η σύγκλιση για να επαναληφθεί πάλι η ίδια διαδικασία. Ετσι υπάρχει η δυνατότητα βελτίωσης της ανακατασκευής, με τη βελτίωση μίας κακής αρχικής λύσης που πιθανόν να οδηγούσε σε συνεχή απόκλιση από την πραγματική λύση με την χρήση μονόδρομης πυραμίδας. Σημαντικός περιορισμός της τεχνικής αυτής είναι το μεγάλο υπολογιστικό κόστος καθώς και η πιθανή παγίδευση του μοντέλου σε διαδοχικές αναστροφές όσον αφορά την κατεύθυνση της επεξεργασίας.

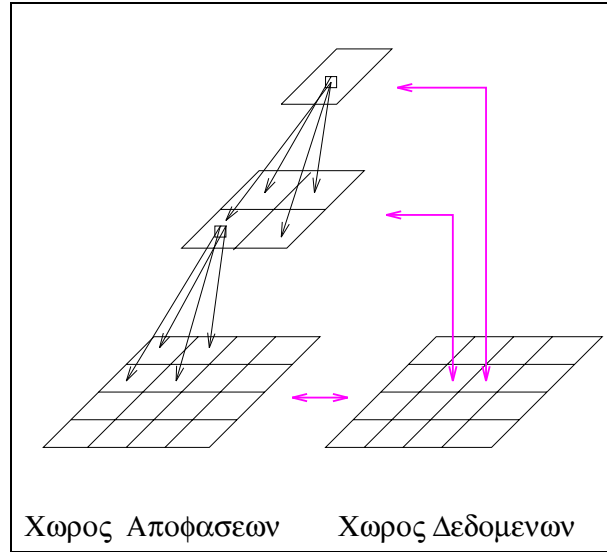
### 2.2.1 Πολλαπλά Επίπεδα στον χώρο Δεδομένων και Αποφάσεων

Η ιδέα στην περίπτωση αυτή είναι αρκετά απλοϊκή και στόχο έχει την δημιουργία ενός απλού μοντέλου ικανού να μας δώσει γρήγορη λύση στο πρόβλημα. Η συνάρτηση κόστους μεταβάλλεται από επίπεδο σε επίπεδο όσον αφορά το πλήθος των όρων που την αποτελούν και όχι την μορφή της η οποία παραμένει η ίδια ([56]).

Ο στόχος είναι η δημιουργία μίας πυραμίδας στο χώρο των παρατηρήσεων (δεδομένων) αλλά και των αποφάσεων. Ετσι χρησιμοποιώντας φίλτρα Gauss, δημιουργείται μια πυραμίδα στο χώρο των παρατηρήσεων, οπότε σε κάθε επίπεδο ελαχιστοποιείται συνάρτηση κόστους στην οποία συμμετέχουν μόνο οι παρατηρήσεις που αντιστοιχούν σε αυτό το επίπεδο και έχουν προκύψει με την χρήση των φίλτρων Gauss (Σχήμα 2.5). Ετσι στο επίπεδο  $i$  η μορφή της συνάρτησης κόστους που ελαχιστοποιείται είναι

$$U^i(\omega^i, d_{S^i}) = \sum_{c \in C_{S^i}} V_c(\omega^i) - T \sum_{s \in S^i} \ln[\delta(\omega_s^i, d_s^i)] \quad (2.8)$$

όπου απαιτείται μία προσαρμογή του όρου  $\delta()$  σε κάθε επίπεδο καθώς και της παραμέτρου  $T$ . Η λύση του προβλήματος προκύπτει με την ελαχιστοποίηση της παραπάνω συνάρτησης κόστους



Σχήμα 2.6: Πολλαπλά Επίπεδα στον Χώρο Αποφάσεων

σε κάθε επίπεδο, δηλαδή

$$\hat{\omega}^i = \operatorname{argmin} \{U^i(\omega^i, d_{S^i})\} \quad (2.9)$$

### 2.2.2 Πολλαπλά Επίπεδα στον Χώρο Αποφάσεων

Ένα πιο ευέλικτο μοντέλο ([13], [50] και [36]) είναι η χρήση του συνόλου των διαθέσιμων παρατηρήσεων σε όλα τα επίπεδα επεξεργασίας. Σαν συνέπεια των παραπάνω, η συνάρτηση κόστους που ορίζει την ανακατασκευή μπορεί να ξαναγραφεί συναρτήσει των αποφάσεων  $\omega \in \Omega_i$  που ορίζονται στο  $S^i$ . Έτσι στο επίπεδο  $i$  μπορούμε να ορίσουμε συνάρτηση κόστους  $U^i(\omega^i, d)$  ισοδύναμη με την  $U(\omega, d)$ , από την οποία προκύπτει η ανακατασκευή  $\omega^i$  στο πλέγμα  $S^i$ ,

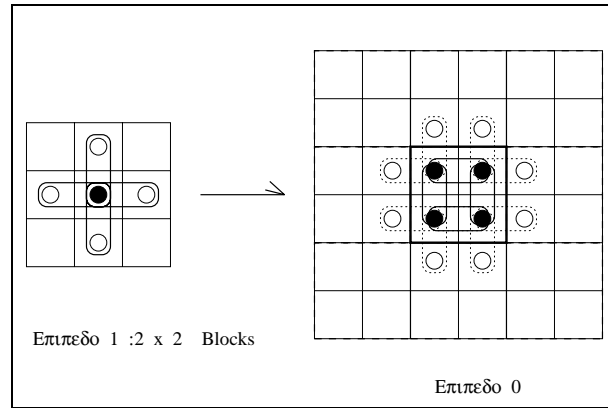
$$U^i(\omega^i, d) = U_1(\Phi^i(\omega^i)) + U_2(\Phi^i(\omega^i), d) \quad (2.10)$$

όπου ο όρος  $\Phi^i$  ορίζει ισομορφικό τελεστή που συσχετίζει την ανακατασκευή  $\omega_i$  στο πλέγμα  $S^i$  με τον πλήρη χώρο παρατηρήσεων και ανακατασκευής.

Είναι εύκολο να δούμε ότι το ούστημα γειτνίασης που ορίζεται αρχικά μπορεί εύκολα να επεκταθεί στα διάφορα επίπεδα, λαμβάνοντας υπόψη τις αλληλεπιδράσεις στο αρχικό επίπεδο. Έτσι μία παρέα δύο σημείων σε κάποιο επίπεδο  $i$  αντιστοιχεί σε παρέες μεταξύ μπλοκ σημείων στο τελευταίο επίπεδο (Σχήμα 2.7) οπότε η μορφή της συνάρτησης κόστους είναι

$$U^i(\omega_i, d) = \sum_{c \in C_{S^i}} \sum_{c' \in C_c} U_s(\omega_i) - T \sum_{s \in S^i} \sum_{s' \in \mathcal{E}_s} \ln[\delta(\omega_s^i, s')] \quad (2.11)$$

Με βάση τους παραπάνω ορισμούς, αντί να ελαχιστοποιούμε την αυθεντική συνάρτηση κόστους πάνω σε πολύ μεγάλο σύνολο λύσεων, χρησιμοποιούμε ένα εναλλακτικό τρόπο



Σχήμα 2.7: Προβολή των σχέσεων παρέας στον πλήρη χώρο δεδομένων

αντιμετώπισης του προβλήματος, όπου υπάρχει η παρακάτω ακολουθία προβλημάτων προς βελτιστοποίηση:

$$\omega(i) = \underset{\omega \in \Omega_i}{\operatorname{argmin}} \{U(\omega, d)\}, i = \operatorname{toplevel}, \dots, 0 \quad (2.12)$$

Ετσι με βάση την εξίσωση (2.10) η παραπάνω σχέση είναι ισοδύναμη με

$$\omega(i) = \underset{\omega^i}{\operatorname{argmin}} \{U^i(\omega^i, d)\}, i = \operatorname{toplevel}, \dots, 0 \quad (2.13)$$

## 2.3 Αιτιοκρατικοί Αλγόριθμοι Χαλάρωσης

Οι αλγόριθμοι που χρησιμοποιούνται για εύρεση ελαχίστων συναρτήσεων κόστους σε προβλήματα ανακατασκευής, πρέπει να εκμεταλλεύονται στο μέγιστο δυνατό τις διαθέσιμες πληροφορίες (παρατηρήσεις) και να έχουν το ελάχιστο δυνατό κόστος. Επίσης πρέπει να πληρούν κάποιες προϋποθέσεις, που είναι:

- **Απόδοση:** Το κόστος των απαιτούμενων υπολογισμών πρέπει να είναι συμβατό με τις απαιτήσεις της εφαρμογής για την οποία χρησιμοποιούνται. Σε ιδανική περίπτωση ένας επαναληπτικός αλγόριθμος δεν πρέπει απλώς να έχει αιτιοκρατική συμπεριφορά, αλλά σε κάθε βήμα να πραγματοποιεί την μέγιστη δυνατή βελτίωση, δηλαδή σε κάθε βήμα να εμπεριέχει τη βέλτιστη λύση του προβλήματος.
- **Προβλεψιμότητα:** Η τελική εκτίμηση πρέπει να εξαρτάται μόνο από τις διαθέσιμες παρατηρήσεις και την εκ των προτέρων κατανομή που περιγράφει το πρόβλημα. Ετσι ο χρήστης δεν πρέπει να είναι υπεύθυνος για την επιλογή παραμέτρων που επηρεάζουν σημαντικά την απόδοση του αλγορίθμου, όπως η αρχική λύση και η διάταξη με την οποία γίνονται οι επισκέψεις στα διάφορα σημεία της εικόνας.

- **Σταθερότητα:** Οι εκτιμήσεις δεν πρέπει να αλλοιώνονται σημαντικά από την παρουσία θορύβου, καθώς και από λάθη στην επιλογή των παραμέτρων του μοντέλου που επιλέγεται για την αναπαράσταση του προβλήματος.

### 2.3.1 Επαναλαμβανόμενες Τοπικά Επικρατούσες Τιμές

Ο αλγόριθμος ETET ([11]) είναι ένα επαναληπτικό αιτιοκρατικό οχήμα χαλάρωσης, ικανό να δώσει λύση σε προβλήματα ελαχιστοποίησης συναρτήσεων κόστους με σχετικά μικρό υπολογιστικό κόστος.

Εστω  $\hat{\omega}$  η προσωρινή λύση του προβλήματος της ανακατασκευής και έστω  $\omega^*$  η λύση που τελικά πρέπει να εξαχθεί. Ο αναμενόμενος στόχος οποιουδήποτε αλγόριθμου, είναι η ενημέρωση της απόφασης  $\hat{\omega}_s$  του σημείου  $s$  χρησιμοποιώντας όλες τις διαθέσιμες παρατηρήσεις, με “αιτιοκρατική” και “λογική” διαδικασία. Η διαδικασία στην οποία στηρίζεται ο προτεινόμενος αλγόριθμος είναι η επιλογή της απόφασης που έχει μέγιστη πιθανότητα, δοθέντων των παρατηρήσεων και της τελευταίας ανακατασκευής  $\hat{\omega}$  για όλα τα σημεία πλην του  $s$ . Έτσι η απόφαση  $\hat{\omega}_s$  που επιλέγεται, δίνει μέγιστη τιμή για την δεσμευμένη πιθανότητα  $p(\omega_s | d, \hat{\omega}_{S/s})$  η οποία είναι ισοδύναμη με την  $p(\omega_s | d, \hat{\omega}_{g_s})$ . Δηλαδή για κάθε σημείο επιλέγεται η απόφαση με την μέγιστη τιμή για την παραπάνω δεσμευμένη πιθανότητα, κάτι που είναι ισοδύναμο με την επιλογή της απόφασης που παρουσιάζει ελάχιστη τοπική ενέργεια ως προς τις παραπάνω παραμέτρους.

Η αρχική ανακατασκευή προκύπτει με την χρήση κριτηρίου μέγιστης πιθανοφάνειας, και στη συνέχεια εφαρμόζεται ο αλγόριθμος μέχρι να συγκλίνει στην τελική ανακατασκευή  $\omega^*$ . Όσον αφορά την δεσμευμένη πιθανότητα ισχύει:

$$p(\omega | d) = p(\omega_s | d, \omega_{g_s}) p(\omega_{g_s} | d) \quad (2.14)$$

Από την παραπάνω σχέση είναι φανερό ότι η  $p(\hat{\omega} | d)$  δεν μειώνεται σε καμιά φάση της επεξεργασίας, γεγονός που μας εξασφαλίζει την σύγκλιση του αλγορίθμου. Από την εφαρμογή του αλγορίθμου παρατηρούμε ότι η σύγκλιση προκύπτει μετά από 5 ή 6 περάσματα.

Με άμεσο στόχο την ανεξαρτησία του αλγορίθμου από την διάταξη που ακολουθείται για τις επισκέψεις στα διάφορα σημεία της εικόνας, στις αποφάσεις στο τρέχον βήμα δεν χρησιμοποιείται η τελευταία ανακατασκευή (συνδυασμός των αποφάσεων που υπάρχουν και αυτών που παίρνονται), αλλά αυτή που έχει προκύψει στο προηγούμενο πέρασμα. Επίσης για αποφυγή διάδοσης λαθών, χρησιμοποιείται ένα επιπλέον ενδεχόμενο στην ανακατασκευή, που είναι ισοδύναμο με “κενή” απόφαση. Έτσι τα σημεία στα οποία το πιο πιθανό ενδεχόμενο παρουσιάζει πολύ μικρή πιθανότητα στην αρχικοποίηση, εξαιρούνται. Για τα σημεία αυτά οι αρχικές αποφάσεις παίρνονται μετά το πρώτο πέρασμα αφού τότε μπορούμε με την χρήση της γειτονιάς να πάρουμε αρκετά πιο σίγουρη απόφαση.

```

Set_Active(S)
 $\omega = \text{Max\_Lik\_Init}(S)$ 
while ( change_pixels  $\neq$  0){
  change_pixels = 0
  for each ( s  $\in$  S )
    If ( state(s) = Active ) {
      Find_Optimal_Label(s)
      If ( Current( $\omega_s$ )  $\neq$  Previous( $\omega_s$ ) ) {
        change_pixels++
        neighbor(s) = Active
      }
    }
  }
}
return(Current( $\omega$ ))

```

Σχήμα 2.8: Επαναλαμβανόμενες Τοπικά Επικρατούσες Τιμές

Τέλος για την μείωση του κόστους δεν επισκεπτόμαστε όλα τα σημεία της εικόνας σε κάθε πέρασμα. Ένα σημείο μπορεί να παρουσιάζει διαφορετικές τιμές στην δεομευμένη πιθανότητα για κάποια απόφαση, μόνο αν υπάρχουν μεταβολές στις αποφάσεις που έχουν ληφθεί για τα γειτονικά του σε σχέση με το προηγούμενο πέρασμα. Έτσι για όλα τα σημεία των οποίων οι αποφάσεις διαφέρουν από αυτές που είχαμε στο προηγούμενο πέρασμα, καθιστούμε τα γειτονικά τους “ενεργά” και τα επισκεπτόμαστε στην συνέχεια (επόμενο πέρασμα). Με την χρήση αυτής της διαδικασίας, αποφεύγονται οι υπολογισμοί τοπικών ενεργειών και πιθανοτήτων που δεν έχουν διαφοροποιηθεί από την τελευταία φορά υπολογισμού τους. Ενεργητική είναι η επίδραση της παραπάνω τεχνικής στα τελευταία περάσματα του αλγορίθμου, όπου ο αριθμός των σημείων που διαφοροποιούνται σε σχέση με τα προηγούμενα περάσματα είναι πολύ μικρός.

### 2.3.2 Μέγιστη Εμπιστοσύνη κατά Προτεραιότητα

Η ιδέα στην οποία στηρίζεται ο αλγόριθμος αυτός ([20]), είναι η βελτιστοποίηση του τρόπου επίσκεψης στα σημεία της εικόνας, με στόχο να παίρνονται αποφάσεις αρχικά για σημεία που έχουν μικρή πιθανότητα να αλλάξουν συμπεριφορά στη συνέχεια. Έτσι αντί να ελαχιστοποιείται

μία ολική ενέργεια, τοπικές ενέργειες που καθορίζουν την σταθερότητα των αποφάσεων υιοθετούνται. Παρατηρούμενα χαρακτηριστικά και εκ των προτέρων γνώση συνδυάζονται για τον υπολογισμό των ενεργειών αυτών. Στην συνέχεια ακολουθεί μια σύντομη και συνοπτική περιγραφή του αλγορίθμου.

Εστω  $L = \{l_1, l_2, \dots, l_Q\}$  όπως έχει ήδη οριστεί το σύνολο των αμοιβαία αποκλειομένων αποφάσεων που αφορούν το πρόβλημα, και  $\Omega$  ο χώρος των πιθανών λύσεων του προβλήματος. Η εκ των υστέρων γνώση σχετικά με τις αποφάσεις αναπαρίσταται με κατανομή Gibbs και η εκ των υστέρων πιθανότητα για κάποιο χάρτη αποφάσεων  $\omega$ , είναι της μορφής:

$$P(\omega|d) = \frac{1}{Z} e^{-\frac{1}{T} \{ \sum_{c \in C} V_c(\omega) + T \sum_{s \in S} \delta(\omega_s | d_s) \}} \quad (2.15)$$

όπου ο όρος  $\delta(\omega_s | d_s)$  εκφράζει την συνέπεια μεταξύ των αποφάσεων ( $\omega$ ) και των διαθέσιμων παρατηρήσεων ( $d_s$ ). Εστω  $\bar{L} = L \cup \{l_0\}$  νέο διευρυμένο σύνολο αποφάσεων, όπου  $l_0$  αντιστοιχεί σε “κενή” απόφαση κατά τη διάρκεια της επεξεργασίας. Εστω επίσης  $\bar{\Omega} = \{\omega = (\omega_1, \dots, \omega_N) | \omega_s \in \bar{L}, \forall s \in S\}$ . Η ιδέα στην οποία βασίζεται ο αλγόριθμος ΜΕκΠ είναι η δημιουργία μίας ακολουθίας ημιτελών ανακατασκευών  $\omega^0, \omega^1, \dots, \omega^f$  του  $\bar{\Omega}$ , αρχίζοντας από την κατάσταση  $\omega^0 = (l_0, \dots, l_0)$  και καταλήγοντας στην κατάσταση  $\omega^f$  που αποτελεί τοπικό ελάχιστο της συνάρτησης κόστους  $U(\omega)$  σε σχέση με το  $\Omega$ .

Ένα σημείο είναι “δεομευμένο” με την απόφαση  $l \in L$  όταν στο βήμα  $t$  της ανακατασκευής ισχύει  $\omega_s^t = l$ , ενώ λέμε ότι είναι “διαθέσιμο” αν  $\omega_s^t = l_0$ . Επίσης κάποιο σημείο που είναι “δεομευμένο” δεν μπορεί σε καμιά περίπτωση να επιστρέψει στην κατάσταση “διαθέσιμο”, χωρίς αυτό να σημαίνει ότι δεν μπορεί να αλλάξει η απόφαση που έχει ληφθεί για το δεομευμένο σημείο.

Ας ορίσουμε τώρα την διευρυμένη τοπική ενέργεια για την απόφαση  $u \in L$  στο σημείο  $s \in S$ , συναρτήσει μιας ολικής τμηματοποίησης (ανακατασκευής)  $\omega' \in \bar{\Omega}$  σαν

$$U_l(u) = \sum_{c \in C_s} V_c(\omega') + \delta(d_s | u) \quad (2.16)$$

όπου  $\omega'$  είναι η ίδια ανακατασκευή με την  $\omega$  εκτός του σημείου  $s$  όπου  $\omega'_s = u$  και

$$V_c(u, \omega_r) = \begin{cases} 0 & , \text{αν } \omega_r = l_0 \\ V_c(u, \omega_r) & , \text{αν } \omega_r \neq l_0 \end{cases} \quad (2.17)$$

για κάθε σημείο  $r$  στην γειτονιά του  $s$ . Η τιμή της  $U_l(u)$  αποτελεί μέτρο εμπιστοσύνης της απόφασης  $u$  για το σημείο  $s$  σε σχέση με τις αποφάσεις που έχουν ληφθεί στην γειτονιά του. Θα πρέπει να σημειώσουμε εδώ ότι μόνο “δεομευμένα” γειτονικά σημεία συμμετέχουν στον υπολογισμό αυτής της ποσότητας, οπότε τα “διαθέσιμα” δεν επηρεάζουν τα υπόλοιπα κατά την διάρκεια των αποφάσεων. Πάντως σε κάθε περίπτωση ένα “διαθέσιμο” σημείο λαμβάνει υπόψη του την κατάσταση που υπάρχει στα γειτονικά ενεργά σημεία πριν αλλάξει κατάσταση.

Με στόχο την βέλτιστη δυνατή ποιότητα όσον αφορά την ανακατασκευή ( $\omega^f$ ), ο κανόνας βάση του οποίου αποφασίζεται η διάταξη επίσκεψης στα σημεία της εικόνας είναι ο εξής:

**Σε κάθε βήμα της ανακατασκευής μόνο στο πιο “ασταθές” σημείο (με βάση την απόφαση του) επιτρέπεται να αλλάξει απόφαση ή να περάσει από την κατάσταση “διαθέσιμο” στην κατάσταση “δεσμευμένο”.**

Η σταθερότητα του σημείου  $s$ , με βάση τον χάρτη αποφάσεων  $\omega$  καθορίζεται ως εξής:

$$G_s(\omega) = \begin{cases} \min_{k \in L, k \neq \omega_s} \{\Delta U_l(k, \omega_s)\} & , \text{αν } \omega_s \in L \\ -\min_{k \in L, k \neq j} \{\Delta U_l(k, j)\} & , \text{αν } \omega_s = l_0 \end{cases} \quad (2.18)$$

όπου

$$\begin{aligned} \Delta U_l(k, j) &= U_l(k) - U_l(j) \\ j \in L : U_l(j) &= \min_{k \in L} \{U_l(k)\} \end{aligned} \quad (2.19)$$

Με βάση τους παραπάνω ορισμούς η αξιοπιστία (σταθερότητα) μιας απόφασης καθορίζεται σαν συνδυασμός της εκ των προτέρων γνώσης του μοντέλου, και των αποφάσεων που λαμβάνονται από βήμα σε βήμα. Μία αρνητική τιμή για το  $G$ , ισοδυναμεί με την ύπαρξη πιο σταθερής κατάστασης (μικρότερη τοπική ενέργεια), κάτω από ένα διαφορετικό χάρτη αποφάσεων. Με βάση τον ορισμό του  $G$ , τα “διαθέσιμα” σημεία έχουν αρνητική τιμή για αυτή την ποσότητα. Η τιμή του  $G$  αντιστοιχεί στην ποσότητα της ενέργειας που θα κερδηθεί ή θα χαθεί αλλάζοντας την τρέχουσα απόφαση σε κάποιο σημείο. Όσο μικρότερη τιμή έχει (αρνητική), τόσο πιο βέβαιη είναι η απόφαση για αλλαγή κατάστασης. Σημαντικό στοιχείο για την κατανόηση του αλγορίθμου αποτελεί ο ρόλος της ποσότητας  $G$ . Ετσι σε κάθε βήμα του αλγορίθμου το σημείο για το οποίο παίρνεται απόφαση, δεν είναι αυτό που παρουσιάζει την μικρότερη τοπική ενέργεια αλλά είναι, είτε αυτό που παρουσιάζει την μέγιστη αρνητική απόσταση μεταξύ των δύο πιο πιθανών αποφάσεων (αν είναι “διαθέσιμο”), είτε αυτό που παρουσιάζει την μέγιστη αρνητική απόσταση μεταξύ της πιο πιθανής απόφασης και της απόφασης που έχει ήδη. Ο αλγόριθμος ΜΕκΠ μπορεί να υλοποιηθεί είτε σειριακά είτε παράλληλα. Στην σειριακή υλοποίηση ένας σωρός (συνά προτεραιότητας) καθορίζει την σειρά επίσκεψης στα διάφορα σημεία της εικόνας, ο οποίος δημιουργείται και τροποποιείται σύμφωνα με τις τιμές της ποσότητας  $G_s$  με τέτοιο τρόπο ώστε στην κορυφή του σωρού να είναι πάντα το σημείο με την μικρότερη τιμή για το  $G$  (πιο ασταθές σημείο). Αλλάζοντας την απόφαση για το σημείο  $s$  επηρεάζονται άμεσα οι τιμές του  $G$  στα γειτονικά του και επομένως και η δομή του σωρού.

Με βάση αυτή την προσέγγιση, για τα σημεία που οι διαθέσιμες πληροφορίες δεν επαρκούν για να μας οδηγήσουν σε μια αρκετά οίγουρη κατάσταση, οι αποφάσεις καθυστερούνται μέχρις



```

 $\omega = (l_0, \dots, l_0)$ 
top = Create_Heap( $\omega$ )
while (G_top < 0){
    s = top;
    Change_State( $\omega_s$ )
    Update_G( $G_s$ )
    Adjust_Heap(s)
    for each ( r  $\in N_s$  ) {
        Update_G( $G_r$ )
        Adjust_Heap(r)
    }
}
return( $\omega$ )

```

Σχήμα 2.9: Μέγιστης Εμπιστοσύνη κατά Προτεραιότητα

ότου με την βοήθεια των γειτονικών σημείων να προκύψει οίγουρη απόφαση. Η τεχνική αυτή καθυστερήσης των αποφάσεων για σημεία με λίγες διαθέσιμες παρατηρήσεις έχει τα εξής πλεονεκτήματα:

- Ελαχιστοποιείται η πιθανότητα αλλαγής προηγούμενων αποφάσεων. Έτσι τα σημεία με ελάχιστη διαθέσιμη πληροφορία, που είναι πολύ πιθανό να τους αποδοθεί λάθος απόφαση (αν αυτό γίνει άμεσα), πρέπει να δεσμευθούν αργότερα όταν η συνεισφορά των γειτονικών σημείων θα είναι ικανή (σε σχέση με τις διαθέσιμες πληροφορίες) να οδηγήσει σε σταθερή απόφαση.
- Μειώνεται σημαντικά η πιθανότητα λάθος πληροφόρησης (συνεισφοράς) άλλων σημείων από κάποιο που “δεσμεύεται”. Ένα σημείο μπορεί να οδηγηθεί σε καλύτερη απόφαση χωρίς λάθος βοήθεια από τα γειτονικά του.

Τέλος όσον αφορά την πολυπλοκότητα του αλγορίθμου, μπορούμε να πούμε ότι η σύγκλιση του πραγματοποιείται σε πεπερασμένο χρόνο. Έτσι απαιτούνται  $O(N)$  συγκρίσεις για την δημιουργία αρχικά του σωρού και  $O(\log_2(N))$  για την διαχείριση του σωρού μετά επίσκεψη, πολλαπλασιαζόμενο με το μέγεθος της γειτονιάς, σε ένα μικρό ποσοστό σημείων σε σχέση με το συνολικό. Το επιπλέον κόστος διαχείρισης του σωρού αντισταθμίζεται αφού κατά την διάρκεια της επεξεργασίας μειώνεται σημαντικά αυτός ο παράγοντας.



## Κεφάλαιο 3

# Εκτίμηση Παραμέτρου Εξομάλυνσης

### 3.1 Εισαγωγή-Ανασκόπηση

Η τεχνική της εξομάλυνσης είναι μία τεχνική που χρησιμοποιείται ευρύτατα σε προβλήματα ανακατασκευής εικόνων ([4], [9] και [58]). Σε ένα τέτοιο πλαίσιο, η εκτίμηση ενός  $N$ -διάστατου διανύσματος  $\omega$  (όπου  $N$  ο αριθμός των σημείων της εικόνας), που συμβολίζει τον χάρτη αποφάσεων προκύπτει από την ελαχιστοποίηση συναρτήσεων κόστους της μορφής

$$U(\omega, d) = \lambda U_1(\omega) + U_2(\omega, d) \quad (3.1)$$

Στην προηγούμενη συνάρτηση κόστους,  $d$  είναι οι διαθέσιμες παρατηρήσεις για την επίλυση του προβλήματος, ενώ  $\omega$  όπως έχει ήδη αναφερθεί ο χάρτης αποφάσεων που πρέπει να ανακατασκευαστεί. Ο όρος  $U_2(\omega, d)$  δίνει ένα μέτρο για την ακρίβεια της ανακατασκευής καθώς το  $\omega$  μεταβάλλεται σε σχέση με τις διαθέσιμες παρατηρήσεις. Ο ρόλος του όρου  $U_1(\omega)$  είναι να εισάγει την υπόθεση για ομοιογενή ανακατασκευή (για γειτονικά σημεία πρέπει να παίρνονται όμοιες αποφάσεις), ενώ η παράμετρος  $\lambda (> 0)$  ελέγχει το μέτρο της ομοιομορφίας όσον αφορά την ανακατασκευή. Επίσης συνεισφέρει στην εξισορρόπηση της συμμετοχής των δύο όρων της συνάρτησης κόστους. Κατά την διάρκεια των τελευταίων χρόνων αρκετοί ερευνητές έχουν ασχοληθεί με τον εφαρμογή της μεθόδου σε προβλήματα ανακατασκευής, και όχι μόνο. Ετσι το θέμα επιλογής βέλτιστης τιμής για την παράμετρο αυτή, παρουσιάζει ιδιαίτερο ενδιαφέρον και παραμένει ανοικτό. Επίσης έχει μελετηθεί και η επίδραση της τιμής αυτής στην ανακατασκευή ([27]). Κοινό συμπέρασμα πολλών διαφορετικών προσεγγίσεων είναι ότι η τιμή της παραμέτρου  $\lambda$  επιδρά σημαντικά στην ανακατασκευή.

Αρκετές προσεγγίσεις έχουν προταθεί για την εκτίμηση αυτής της παραμέτρου. Απλή λύση είναι η επιλογή βασισμένη σε στατιστικά μοντέλα, όπου ένα κριτήριο ψευδο-πιθανοφάνειας

του πεδίου Gibbs δίνει την βέλτιστη λύση, είτε με την χρήση αλγόριθμου Monte-Carlo, είτε με τον υπολογισμό ροπών της κατανομής που περιγράφει την ψευδο-πιθανοφάνεια ([25]). Επίσης ενδιαφέρουσα αντιμετώπιση του προβλήματος είναι και με την χρήση της μεθόδου cross-validation. Στην περίπτωση αυτή ορίζεται μία απόσταση μεταξύ των αποφάσεων και των διαθέσιμων παρατηρήσεων, και η βέλτιστη τιμή του  $\lambda$  προκύπτει από την ελαχιστοποίηση μίας συνάρτησης αυτής της απόστασης ([53] και [55]). Τέλος πιο γενική μέθοδος είναι η από κοινού ανακατασκευή και εκτίμηση αυτής της παραμέτρου ([33] και [44]). Η μέθοδος αυτή βασίζεται στις δεομευμένες πυκνότητες πιθανότητας που ακολουθούν οι κατανομές των παρατηρήσεων δεδομένου του χάρτη αποφάσεων  $\omega$ .

Αρκετά ενδιαφέρουσα στατιστική μοντελοποίηση του προβλήματος έγινε από τους A. Γαλατσάνο και N. Κατοάγγελο ([33]), όπου ειδικές υποθέσεις γίνονται όσον αφορά την συσχέτιση μεταξύ των διαθέσιμων παρατηρήσεων και του χάρτη αποφάσεων. Το πρόβλημα είναι κακώς τιθέμενο, δηλαδή οι διαθέσιμες παρατηρήσεις δεν μπορούν να μας εξασφαλίσουν μία και μοναδική λύση, και ειδικές υποθέσεις γίνονται για τον περιορισμό των δυνατών λύσεων και την εξαγωγή της βέλτιστης. Η τιμή της παραμέτρου συσχετίζεται άμεσα με την διασπορά του θορύβου στα δεδομένα. Η ανακατασκευή γίνεται με χρήση κριτηρίου ανάλυσης λάθους, το οποίο βασίζεται σε ελαχιστοποίηση μέσου τετραγωνικού σφάλματος. Επιπλέον προτείνεται ένα εναλλακτικό σχήμα για την εκτίμηση του  $\lambda$ , καθώς και τρόπος υπολογισμού της διασποράς του θορύβου στα δεδομένα. Η επιλογή της μεθόδου που θα χρησιμοποιηθεί, βασίζεται στην γνώση για το μοντέλο και την μορφή του θορύβου. Παρόμοια προσέγγιση προτείνεται στο ([27]), όπου η επιλογή του  $\lambda$  γίνεται ανάλογα με την ισχύ του θορύβου. Διαχωρισμός μεταξύ των περιπτώσεων χαμηλής και υψηλής ισχύος στο θόρυβο λαμβάνει χώρα. Ετσι για περιπτώσεις θορύβου χαμηλής ισχύος στατιστικά κριτήρια χρησιμοποιούνται για την επιλογή του  $\lambda$ , ενώ η χρήση cross-validation τεχνικών υιοθετείται σε περιπτώσεις με ισχυρό θόρυβο. Τέλος υπάρχουν ανακεφαλαιωτικά συμπεράσματα και αξιολόγηση των παραπάνω μεθόδων όσον αφορά τα πλεονεκτήματα και τα μειονεκτήματά τους ([5] και [59]).

Συνήθως στις μεθόδους αυτές τα προβλήματα της ανακατασκευής και της βελτιστοποίησης του  $\lambda$  αντιμετωπίζονται ενιαία, κάτι που θεωρητικά οδηγεί σε πιο σταθερή λύση του προβλήματος. Επίσης δεν υπάρχουν περιορισμοί όσον αφορά το είδος του ανακατασκευαζόμενου μοντέλου. Το κύριο μειονέκτημα των μεθόδων αυτών είναι το μεγάλο υπολογιστικό κόστος που φαίνεται να υπάρχει. Αρκετά πολύπλοκες μαθηματικές διεργασίες απαιτούνται (όπως αντιστροφές, ψευδοαντιστροφές, κ.τ.λ) σε αρκετά μεγάλους πίνακες (της τάξης του πλήθους των σημείων της εικόνας). Επίσης ειδικές υποθέσεις γίνονται όσον αφορά την δομή και την μορφή αυτών των πινάκων. Τέλος σε αρκετές περιπτώσεις απαιτείται η γνώση του μοντέλου του θορύβου, κάτι που δεν είναι εύκολο σε πραγματικές εφαρμογές.

### 3.2 Εκτίμηση της παραμέτρου εξομάλυνσης με Ψηφοδοσία

Στην εργασία αυτή το πρόβλημα αντιμετωπίζεται κάτω από διαφορετική οπτική γωνία. Τα δύο προβλήματα διαχωρίζονται όσον αφορά την λύση τους. Η λύση τους προκύπτει εναλλακτικά, και η γενική ιδέα είναι να χρησιμοποιηθεί η τελευταία λύση (για όποιο υπάρχει) του ενός, ώστε να προκύψει μία πιο αξιόπιστη λύση για το άλλο. Έτσι χρησιμοποιώντας τον χάρτη αποφάσεων (ή ένα τμήμα του), είναι δυνατή η εκτίμηση πιο αξιόπιστης τιμής για το  $\lambda$ , η οποία στην συνέχεια θα συμβάλει στην δημιουργία ενός νέου χάρτη αποφάσεων. Σαν πιο αξιόπιστη τιμή του  $\lambda$  επιλέγεται αυτή που επιβεβαιώνει με τον καλύτερο δυνατό τρόπο τον τελευταίο χάρτη αποφάσεων. Η προτεινόμενη μέθοδος μπορεί να χρησιμοποιηθεί σε προβλήματα ανακατασκευής με πεπερασμένο αριθμό αποφάσεων, όσον αφορά το ανακατασκευαζόμενο σχήμα.

Ας προχωρήσουμε τώρα σε πιο σαφή ορισμό του προβλήματος. Όπως ήδη έχει αναφερθεί ο ρόλος τού όρου  $U_1(\omega)$  είναι να εισάγει την υπόθεση για συμβατές (ίδιες) αποφάσεις σε γειτονικά σημεία. Η χρήση σχέσεων γειννίας (Σχήμα 2.1) καθώς και της μεθόδου εξομάλυνσης συνδέεται άμεσα με την χρήση πεδίων Markov σε προβλήματα ανακατασκευής, όπου οι χάρτες αποφάσεων μοντελοποιούνται με κατανομές Gibbs. Σε μια τέτοια περίπτωση οι χάρτες αποφάσεων προκύπτουν με ελαχιστοποίηση ολικών συναρτήσεων κόστους (μορφή όμοια με αυτήν της εξίσωσης (3.1)). Εστω επίσης ότι ένα τμήμα του χάρτη αποφάσεων είναι ήδη γνωστό, και  $\omega_s$  η απόφαση που έχει αποδοθεί στο σημείο  $s$  σύμφωνα με αυτό τον χάρτη αποφάσεων. Με όμοιο τρόπο με αυτόν που χρησιμοποιείται για τον ορισμό της ολικής συνάρτησης κόστους, μπορούν να οριστούν και τοπικές συναρτήσεις κόστους, αποτελώντας προβολή της ολικής συνάρτησης κόστους σε ένα σημείο και στην γειτονιά του. Εστω  $U_l(\omega, \omega_{g_s}, d_s)$  η τοπική συνάρτηση κόστους στο σημείο  $s$ , για την απόφαση  $\omega$ , δεδομένου των διαθέσιμων παρατηρήσεων  $d_s$  που σχετίζονται με το σημείο  $s$  και του χάρτη αποφάσεων  $\omega_{g_s}$  στην γειτονιά του  $s$ . Συνήθως η γενική μορφή που ακολουθεί η συνάρτηση αυτή είναι:

$$U_l(\omega, \omega_{g_s}, d_s) = \delta(\omega, d_s) + \lambda \sum_{c \in C_s} V_c(\omega) \quad (3.2)$$

όπου ο όρος  $C_s$  συμβολίζει τις παρέες στην γειτονιά του  $s$ ,  $g_s$ , και  $V_c(\omega)$  συνάρτηση δυναμικού,

$$V_c(\omega_s, \omega_t) = \begin{cases} -1, & \text{αν } \omega_s = \omega_t \\ +1, & \text{αν } \omega_s \neq \omega_t \end{cases} \quad (3.3)$$

Χωρίς βλάβη της γενικότητας και για λόγους απλότητας μπορούμε να υποθέσουμε ένα σύνολο αποφάσεων με δύο πιθανά γεγονότα, όσον αφορά την ανακατασκευή. Εστω  $p(\omega_s | \omega_{g_s}, d_s, \lambda)$  η δεομευμένη πυκνότητα πιθανότητας για την απόφαση  $\omega_s$  στο σημείο  $s$ . Σε μια τέτοια περίπτωση η τιμή της συνάρτησης πυκνότητας πιθανότητας για την τρέχουσα απόφαση πρέπει να είναι μεγαλύτερη, από την αντίστοιχη τιμή της για την αντίθετη απόφαση, δηλαδή η παρακάτω

ανισότητα πρέπει να ισχύει:

$$p(\omega_s | \omega_{g_s}, d_s, \lambda) \geq p(\bar{\omega}_s | \omega_{g_s}, d_s, \lambda) \quad (3.4)$$

όπου  $\bar{\omega}_s$  η αντίθετη απόφαση της  $\omega_s$ . Το πρόβλημα στην παραπάνω ανισότητα είναι ότι τα μεγέθη που εμφανίζονται δεν είναι άμεσα μετρήσιμα. Για το λόγο αυτό ανάγουμε την παραπάνω ανισότητα σε ανισότητα με όρους τις τοπικές συναρτήσεις κόστους (ενέργειες), οπότε προκύπτει νέα μορφή πιο ευέλικτη αποτελούμενη από όρους άμεσα μετρήσιμους, δηλαδή

$$U_l(\omega_s, \omega_{g_s}, d_s) \leq U_l(\bar{\omega}_s, \omega_{g_s}, d_s) \quad (3.5)$$

Εστω  $N_{total}$  ο μέγιστος αριθμών παρεών στην χειτονιά του  $s$  που είναι ενεργές (στα σημεία που συμμετέχουν έχουν αποδοθεί αποφάσεις), και έστω  $n_{\omega_s}$  ο αριθμός των παρεών με αρνητική συνεισφορά στην τοπική συνάρτηση κόστους (την μειώνουν). Χρησιμοποιώντας τις παραπάνω υποθέσεις η σχέση (3.2) γράφεται:

$$U_l(\omega_s, \omega_{g_s}, d_s) = \delta(\omega_s, d_s) + \lambda(N_{total} - 2n_{\omega_s}) \quad (3.6)$$

Αντικαθιστώντας την νέα μορφή της ενέργειας  $U_l$  στην σχέση (3.5) έχουμε

$$\delta(\omega_s, d_s) + \lambda(N_{total} - 2n_{\omega_s}) \leq \delta(\bar{\omega}_s, d_s) + \lambda(N_{total} - 2n_{\bar{\omega}_s}) \quad (3.7)$$

όπου  $n_{\bar{\omega}_s}$  ο αριθμός των παρεών με θετική συνεισφορά στην τοπική ενέργεια κόστους  $U_l$  (την αυξάνουν). Αφού  $n_{\bar{\omega}_s} = N_{total} - n_{\omega_s}$  η προηγούμενη σχέση δίνει

$$\delta(\omega_s, d_s) - \delta(\bar{\omega}_s, d_s) + 2\lambda(N_{total} - 2n_{\omega_s}) \leq 0 \quad (3.8)$$

Από την παραπάνω ανισότητα μπορούν να εξαχθούν περιορισμοί όσον αφορά την τιμή του  $\lambda$ . Αν χρησιμοποιήσουμε τις ήδη γνωστές τιμές που υπάρχουν για τους όρους  $\omega_s$ ,  $\bar{\omega}_s$  και  $d_s$  η μόνη άγνωστη παράμετρος είναι το  $\lambda$ . Παρόμοια ανισότητα μπορεί να εξαχθεί για κάθε σημείο της εικόνας που υπάρχει ήδη απόφαση, ορίζοντας μία ημιευθεία. Θεωρητικά η τομή των περιοχών που ορίζονται από τις παραπάνω ανισότητες μπορεί να μας δώσει την βέλτιστη τιμή του  $\lambda$ , αλλά το πρόβλημα είναι ότι κάτι τέτοιο απαιτεί μεγάλο υπολογιστικό κόστος και επίσης η επίδραση του θορύβου μπορεί να οδηγήσει σε κενό σύνολο. Για να αντιμετωπιστούν τα προβλήματα αυτά ο χώρος αναζήτησης για το  $\lambda$  κβαντίζεται, δηλαδή συνεχή διαστήματα αντιστοιχίζονται με μία κβαντισμένη τιμή. Ανάλογα με την μορφή της σχέσης (3.8), υπάρχουν τιμές του  $\lambda$  που την επαληθεύουν. Οι τιμές αυτές λαμβάνουν ψήφο, η οποία δεν είναι δυαδική, αλλά έχει βάρος το οποίο εξαρτάται από την συνεισφορά της δοσμένης τιμής του  $\lambda$ , στην ελαχιστοποίηση της συνάρτησης κόστους. Δηλαδή οι ψήφοι δίνονται έτσι ώστε να επιβεβαιωθεί με τον καλύτερο δυνατό τρόπο ο χάρτης αποφάσεων.

Συνηθισμένο πρόβλημα στην ανακατασκευή εικόνων με την χρήση μεθόδου εξομάλυνσης, είναι η υπερβολική λείανση του ανακατασκευαζόμενου πεδίου. Αυτό γίνεται γιατί συνήθως

```

create_ML_labeling( $\omega$ );
while ( 1 ){
  for each ( site ) {
    for each (  $\lambda$  )
      if (  $energy(\omega_s) < energy(\overline{\omega}_s)$  )
        give_vote( $\lambda$ );
     $\lambda = best\_votes\_selection()$ ;
     $search\_space\_center = \lambda$ ;
     $sub\_divide(\lambda_{distance})$ 
     $\omega = create\_new\_labeling(\lambda)$ ;
  }
return( $\omega$ )

```

Σχήμα 3.1: Εκτίμηση Παραμέτρου Εξομάλυνσης με Ψηφοδοσία

μεγάλες τιμές του  $\lambda$ , έχουν σημαντική συνεισφορά στην ελαχιστοποίηση της συνάρτησης κόστους και τείνουν να εξαλείψουν την παρουσία και την συμμετοχή του όρου  $U_2(\omega, d)$ , που εκφράζει την συμβατότητα των αποφάσεων με τις διαθέσιμες παρατηρήσεις. Παρόμοια προβλήματα υπάρχουν και στην προτεινόμενη λύση, γεγονός που επιβάλλει την τροποποίηση των βαρών των ψήφων, ώστε να λαμβάνεται υπόψη και αυτό το ενδεχόμενο. Για τον λόγο αυτό τα βάρη στις ψήφους δίνονται με τέτοιο τρόπο, ώστε αφενός να επιβεβαιώνουν με τον καλύτερο δυνατό τρόπο τον τελευταίο χάρτη αποφάσεων, και αφετέρου να μην προκαλούν υπερβολική λείανση του ανακατασκευαζόμενου πεδίου. Όσον αφορά το πρώτο στόχο, κριτήριο αποτελεί η διαφορά μεταξύ των τοπικών ενεργειών για την τρέχουσα απόφαση και την αντίθετη της. Όσο μεγαλύτερη είναι αυτή η διαφορά (η οποία εξαρτάται άμεσα από την τιμή του  $\lambda$ ), τόσο πιο αποδεκτή είναι η δοσμένη τιμή του  $\lambda$  (με την προϋπόθεση φυσικά ότι η ενέργεια για την τρέχουσα απόφαση είναι μικρότερη). Εφόσον υπάρχει ήδη ένα διαθέσιμο ανακατασκευαζόμενο τμήμα, μπορεί πάνω σε αυτό να υπολογιστεί η μέση τιμή καθώς και η διασπορά του όρου  $\delta(\omega_s, d_s)$ . Η μέση τιμή εκφράζει την συμμετοχή του όρου αυτού στην ελαχιστοποίηση σε τοπικό επίπεδο, οπότε μπορεί να εξισορροπηθεί με την συμμετοχή του άλλου όρου,

$$\lambda \sum_{c \in C_s} V_c(\omega) \quad (3.9)$$

ανάλογα με την απόστασή τους. Όσο πιο κοντά είναι οι τιμές των δύο αυτών όρων τόσο πιο σίγουρη και σταθερή είναι η δοσμένη τιμή του  $\lambda$ . Επίσης η διασπορά του όρου μας δίνει κάποια ένδειξη όσον αφορά και τις πιθανές διακυμάνσεις που μπορεί να έχει ο όρος  $\delta(\omega_s, d_s)$ , και

πρέπει επίσης να ληφθεί υπόψη. Με βάση τις προηγούμενες παρατηρήσεις το βάρος της ψήφου για τις διάφορες τιμές του  $\lambda$ , που ικανοποιούν την ανισότητα (3.8) είναι,

$$\beta(\lambda|\omega_s) = \frac{U_l(\bar{\omega}_s, \omega_{g_s}, d_s) - U_l(\omega_s, \omega_{g_s}, d_s)}{\sigma^2(\delta(w, d)) + (\lambda \sum_{c \in C_s} V_c(\omega) + E(\delta(w, d)))^2} \quad (3.10)$$

Με την μέθοδο αυτή είναι εφικτός ο προσδιορισμός μιας τιμής για το  $\lambda$ . Η τιμή αυτή χρησιμοποιείται και ένας νέος χάρτης αποφάσεων προκύπτει. Πριν ξαναχρησιμοποιηθεί η μέθοδος για προσδιορισμό νέας τιμής του  $\lambda$ , η ήδη υπάρχουσα τιμή γίνεται κέντρο του χώρου αναζήτησης, ενώ ταυτόχρονα η απόσταση μεταξύ δύο διαδοχικών κβαντισμένων τιμών καθώς και ο χώρος αναζήτησης, υποδιπλασιάζονται. Έτσι αφενός υπάρχει μεγαλύτερη ακρίβεια στον προσδιορισμό του  $\lambda$  (αφού οι αποστάσεις μεταξύ δύο διαδοχικών τιμών είναι μικρότερη) και αφετέρου η αναζήτηση περιορίζεται σε χώρο που είναι πιο πιθανή η εξαγωγή βέλτιστης τιμής.

Η μέθοδος αυτή μπορεί πολύ εύκολα να επεκταθεί και σε εφαρμογές με περισσότερες των δύο δυνατές αποφάσεις. Στο ενδεχόμενο αυτό οι ενεργοί περιορισμοί για το  $\lambda$  προκύπτουν από τομή ανισοτήτων της μορφής (3.8) και πιο συγκεκριμένα ισχύει:

$$\bigcap_{\substack{\forall u \in L \\ u \neq \omega_s}} (U_l(\omega_s, \omega_{g_s}, d_s) < U_l(u, \omega_{g_s}, d_s)) \quad (3.11)$$

όπου  $L$  το σύνολο αποφάσεων, ενώ για τον υπολογισμό των βαρών των ψήφων οι διαφορές μεταξύ των ενεργειών της τρέχουσας απόφασης και των υπολοίπων λαμβάνονται υπόψη. Τέλος η μέθοδος μπορεί να επεκταθεί και σε εφαρμογές με περισσότερες από μία παραμέτρους εξομάλυνσης. Οι ψήφοι δεν δίνονται πλέον σε μεμονωμένες τιμές αλλά σε διανύσματα τιμών. Έτσι χρησιμοποιούνται όλοι οι δυνατοί συνδυασμοί μεταξύ των κβαντισμένων τιμών των παραμέτρων εξομάλυνσης σαν πιθανά βέλτιστα διανύσματα τιμών. Το πρόβλημα του μεγάλου υπολογιστικού κόστους σε αυτή την περίπτωση (λόγω του μεγάλου αριθμού πιθανών συνδυασμών), μπορεί να αντιμετωπιστεί με δραστική μείωση των κβαντισμένων τιμών των διαφόρων παραμέτρων εξομάλυνσης, όχι κατά ανάγκη με όμοιο τρόπο. Έτσι για την πιο σημαντική παράμετρο μπορούμε να κρατήσουμε ικανοποιητικό αριθμό κβαντισμένων τιμών, ενώ για λιγότερο σημαντικές να τον μειώσουμε. Η προτεινόμενη μέθοδος χρησιμοποιήθηκε στο πρόβλημα της ανίχνευσης μεταβολών.



## Κεφάλαιο 4

# Ανίχνευση Κινούμενων Αντικειμένων

### 4.1 Εισαγωγή-Ανασκόπηση

Το πρόβλημα της ανίχνευσης μεταβολών συνίσταται από μία δυαδική απόφαση  $\Theta(i, j)$  για κάθε σημείο  $(i, j)$  στο πλέγμα στο οποίο προβάλλονται οι εικόνες στον διοδιάστατο χώρο. Για το τυχαίο πεδίο  $\Theta(i, j)$  έχουμε δύο πιθανά γεγονότα,  $\Theta(i, j) = s$  (Static), αν το σημείο ανήκει στο υπόστρωμα ( $H_0$ ), και  $\Theta(i, j) = m$  (Mobile), αν το σημείο ανήκει σε ανεξάρτητα κινούμενο αντικείμενο. Με βάση αυτές τις υποθέσεις για κάθε σημείο μπορούμε να γράψουμε

$$\begin{aligned} H_0 : \Theta(i, j) &= s \\ H_1 : \Theta(i, j) &= m \end{aligned} \tag{4.1}$$

Η λύση του προβλήματος αυτού αποτελεί αρχικό βήμα σε πολλές εφαρμογές της Μηχανικής Ορασης και της Ρομποτικής, ιδιαίτερα στον τομέα της Ανάλυσης Κίνησης. Αρκετοί ερευνητές έχουν προσεγγίσει το θέμα κάτω από διαφορετική οπτική γωνία και αρκετές διαφορετικές λύσεις έχουν προταθεί.

Η πιο απλή μέθοδος για ανίχνευση μεταβολών στηρίζεται σε μεθόδους κατώφλιού, όπου χρησιμοποιείται η απόλυτη τιμή της διαφοράς των φωτεινοτήτων. Ετσι τα σημεία όπου η απόλυτη τιμή της είναι πάνω από κάποιο εμπειρικά καθοριζόμενο κατώφλι, θεωρούνται ότι ανήκουν σε κινούμενα αντικείμενα ([26]). Αμεσες επεκτάσεις του μοντέλου κατώφλιού, με στόχο την σταθερότητα έναντι του θορύβου, υιοθετούν διαφορές φωτεινοτήτων ανά σύνολο σημείων (μπλοκ) ([61]). Βελτιώσεις του μοντέλου οδήγησαν σε στατιστική οριοθέτηση, όπου παρατηρούμενη κατανομή της διαφοράς φωτεινοτήτων στο ζεύγος των εικόνων, προσεγγίζεται με μίξη δύο κατανομών ίδιου τύπου (συνήθως κατανομές Laplace ή Gauss χρησιμοποιούνται γι' αυτό τον λόγο) ([60]). Ετσι με κριτήριο την μεγιστοποίηση της εκ των υστέρων πιθανότητας,

ένας δυναμικός καθορισμός του κατωφλιού λαμβάνει χώρα, εξαρτώμενος από τις παραμέτρους των κατανομών που εκφράζουν τις δύο δυνατές καταστάσεις. Το κυριότερο μειονέκτημα αυτών των προσεγγίσεων, είναι η ευπάθεια (ευαισθησία) τους στον θόρυβο. Επίσης η μέθοδος για ανεξάρτητη απόφαση ανά σημείο, μπορεί να οδηγήσει σε μη συνεπή αποτελέσματα, όπου μία διαδικασία τμηματοποίησης απαιτείται για την δημιουργία συνεκτικών περιοχών, οι οποίες εμφανίζονται συνήθως και στον πραγματικό κόσμο.

Μία διαφορετική κλάση αλγορίθμων, βασισμένη σε στατιστική μοντελοποίηση του προβλήματος επίσης διερευνήθηκε. Οι Cafforio και Rocca ([19]) θεώρησαν ότι οι διαφορές των φωτεινοτήτων είναι στατιστικά ανεξάρτητες ανά σημείο και αυτές σε κάθε σημείο ακολουθούν κατανομή Laplace με μηδενική μέση τιμή. Το πεδίο τμηματοποίησης μοντελοποιείται με πρώτης τάξης αλυσίδες Markov, κατά μήκος των γραμμών της εικόνας, και η λύση του προβλήματος προκύπτει με μεγιστοποίηση της εκ των υστέρων πιθανότητας. Επέκταση του μοντέλου αυτού στον διδιάστατο χώρο των αλυσίδων Markov έχει επίσης προταθεί ([29] και [46]). Η διαφορά μεταξύ των [29] και [46] υπάρχει στην κατανομή πυκνότητας πιθανότητας που χρησιμοποιείται για να περιγράψει την διαφορά των φωτεινοτήτων. Κοινό χαρακτηριστικό και των δύο μεθόδων είναι ότι οι πιθανότητες μετάβασης από την μία κατάσταση στην άλλη καθορίζονται εμπειρικά.

Η Μπαγιασιανή μοντελοποίηση (Bayesian formulation) του προβλήματος έχει επίσης διερευνηθεί. Η χρήση πιο πολύπλοκων μοντέλων, όπου τυχαία πεδία Markov (MRFs) με την χρήση κατανομών Gibbs περιγράφουν το πρόβλημα, μπορεί να αυξάνει την πολυπλοκότητα του προβλήματος, αλλά έχει ευεργετική επίδραση στην ποιότητα της ανακατασκευής. Η λύση του προβλήματος βασίζεται στην δημιουργία ολικών συναρτήσεων κόστους, όπου εξαρτήσεις και συσχετίσεις ορίζονται μεταξύ μη γραμμικών μεγεθών και η ανακατασκευή γίνεται με την ελαχιστοποίηση των συναρτήσεων αυτών. Η πιο απλή προσέγγιση υιοθετεί πεδία Markov αλλά και τεχνικές κατωφλιού ([1]). Η τιμή του κατωφλιού καθορίζεται δυναμικά με βάση τις ιδιότητες των πεδίων Markov, διαφέροντας από σημείο σε σημείο ανάλογα με τις τιμές των διαθέσιμων παρατηρήσεων στο δοσμένο σημείο και τις αποφάσεις που έχουν ληφθεί για τα γειτονικά σημεία.

Η χρήση χώρο-χρονικών μοντέλων Markov οδήγησε σε πιο πολύπλοκες προσεγγίσεις, όπου ο χάρτης αποφάσεων προκύπτει από ελαχιστοποίηση συναρτήσεων κόστους με χρήση αιτιοκρατικών αλγορίθμων ([12], [14], [16], [18], [41] και [54]). Η διαφοροποίηση μεταξύ των παραπάνω μεθόδων οφείλεται κυρίως στην μορφή και στον τρόπο κατασκευής των συναρτήσεων κόστους, ενώ και διαφορετικοί επαναληπτικοί αλγόριθμοι χρησιμοποιούνται για την ελαχιστοποίησή της. Ενδεικτικά θα αναφερθούμε σε κάποιες από αυτές.

Κοινό σημείο όλων τα παραπάνω προσεγγίσεων είναι ότι η ελαχιστοποίηση γίνεται χρησιμοποιώντας αιτιοκρατικούς αλγόριθμους χαλάρωσης ([41]), που συνήθως είναι αρκετά ευαίσθητοι στην αρχική λύση του προβλήματος. Ένα πιο πολύπλοκο πρόβλημα, όπου εξετάζεται και η παρουσία υφής (texture), αναλύεται στο ([54]), όπου το σύνολο των πιθανών αποφάσεων

έχει διοδιάστατα διανύσματα,  $\{Static, Mobile\} \times \{Texture, Smooth\}$ . Η χρήση αλγορίθμου πολλαπλών επιπέδων (συνήθως 2) εξασφαλίζει μία γρήγορη επίλυση, αφού και το κόστος του αλγορίθμου ETET που χρησιμοποιείται για την ελαχιστοποίηση σε κάθε επίπεδο είναι αρκετά μικρό. Επίσης αρκετές διαφορετικές τμηματοποιήσεις εξετάζονται σαν αρχική λύση για τον αλγόριθμο, με στόχο την αποφυγή παγίδευσης του σε τοπικά ελάχιστα. Παραμφερής είναι και ο τρόπος αντιμετώπισης του προβλήματος στο [18]. Το πρόβλημα οριοθετείται με χρήση τεχνικής εξομάλυνσης, όπου ιδιαίτερη προσοχή έχει δοθεί στον ορισμό-καθορισμό της συνάρτησης κόστους, η οποία φαίνεται να είναι αρκετά πολύπλοκη και επίσης αρκετά ακριβή (με την έννοια του υπολογιστικού κόστους). Το πρόβλημα του μεγάλου υπολογιστικού κόστους αντιμετωπίζεται επιτυχημένα με χρήση αλγορίθμου πολλαπλών επιπέδων, όπου η συνάρτηση κόστους που ελαχιστοποιείται σε κάθε επίπεδο αποτελεί μετασχηματισμό αυτής που έχουμε στο τελευταίο. Τα πολλαπλά επίπεδα εφαρμόζονται στο μοντέλο Markov, ενώ σε κάθε επίπεδο χρησιμοποιούνται οι παρατηρήσεις σε πλήρη ανάλυση. Επίσης αντιμετωπίζεται με επιτυχία και η περίπτωση κινούμενου παρατηρητή, με χρήση αλγορίθμου εκτίμησης της ίδιας κίνησης.

Ο αλγόριθμος που προτείνεται σε αυτή την εργασία ανήκει στην κατηγορία των τελευταίων, όπου μοντέλα Markov και κατανομές Gibbs χρησιμοποιούνται για να περιγράψουν το πρόβλημα. Η καινοτομία της μεθόδου είναι η μοντελοποίηση της διαφοράς των φωτεινότητων σαν μίξη κατανομών Laplace, γεγονός που επιτρέπει την δημιουργία όρου που εκφράζει σε πολύ ικανοποιητικό βαθμό την συσχέτιση μεταξύ των παρατηρήσεων και των αποφάσεων. Το πρόβλημα ανάγεται σε πρόβλημα εξομάλυνσης, όπου η προτεινόμενη μέθοδος ψηφοδοσίας για την επιλογή της παραμέτρου εξομάλυνσης εφαρμόζεται με πολύ ικανοποιητικά αποτελέσματα. Έτσι δημιουργείται ένα δυναμικό μοντέλο, όπου δεν απαιτείται εκ των προτέρων γνώση, το οποίο μπορεί να αντιμετωπίσει και περιπτώσεις κινούμενου παρατηρητή, αφού υπάρχει η δυνατότητα για εκτίμηση της ίδιας κίνησης.

## 4.2 Εκτίμηση Ιδίας Κίνησης

Μια αρκετά συνηθισμένη υπόθεση σε προβλήματα ανίχνευσης κινούμενων αντικειμένων είναι η ύπαρξη στατικού παρατηρητή. Στην περίπτωση αυτή είναι αναμενόμενο ότι οι μεταβολές που παρατηρούνται στις φωτεινότητες μεταξύ του ζεύγους των εικόνων να οφείλονται σε κινούμενα αντικείμενα. Η υπόθεση αυτή δεν ανταποκρίνεται στα δεδομένα του πραγματικού κόσμου, όπου συνήθως κινούμενα οπτικά συστήματα έχουν να αντιμετωπίσουν το πρόβλημα της ανίχνευσης άλλων κινούμενων αντικειμένων.

Το πρόβλημα της εκτίμησης της ίδιας κίνησης (κίνηση του παρατηρητή) αντιμετωπίζεται στην εργασία αυτή σε αρχικό στάδιο. Στην συνέχεια αφού η ίδια κίνηση εκτιμηθεί και αφαιρεθεί από την ακολουθία εισόδου, αντιμετωπίζεται το πρόβλημα της ανίχνευσης μεταβολών.

Για την εκτίμηση αυτής της κίνησης χρησιμοποιείται ένας απλός και γρήγορος αλγόριθμος πολλαπλών επιπέδων ([23] και [47]). Η κίνηση αυτή περιγράφεται στο διοδιάστατο χώρο (στον χώρο προβολής), με γραμμικό μοντέλο 6-παραμέτρων, γνωστότερο ως αφφινικό, και πιο συγκεκριμένα ισχύει

$$\begin{pmatrix} u(i, j) \\ v(i, j) \end{pmatrix} = \begin{pmatrix} a_{1,1} + a_{1,2}i + a_{1,3}j \\ a_{2,1} + a_{2,2}i + a_{2,3}j \end{pmatrix} \quad (4.2)$$

Το αφφινικό μοντέλο που χρησιμοποιείται δεν μας εξασφαλίζει τον σωστό υπολογισμό της ίδιας κίνησης, αφού υπάρχουν περιπτώσεις τριοδιάστατης κίνησης που δεν μπορούν να αναπαρασταθούν στο 2-Δ χώρο με αυτό το μοντέλο. Έτσι η χρήση του ενδεικνύεται για τις περιπτώσεις που η κίνηση είναι μεταφορική και όλα τα σημεία της 3-Δ σκηνής βρίσκονται στο ίδιο βάθος, ή σε πολύ μεγάλο βάθος (γεγονός που ισοδυναμεί με παράλληλη προβολή). Επιπλέον μπορεί να χρησιμοποιηθεί και στην περίπτωση μόνο περιστροφικής κίνησης ως προς τον άξονα  $Z$  (περιστροφή  $\gamma$ ).

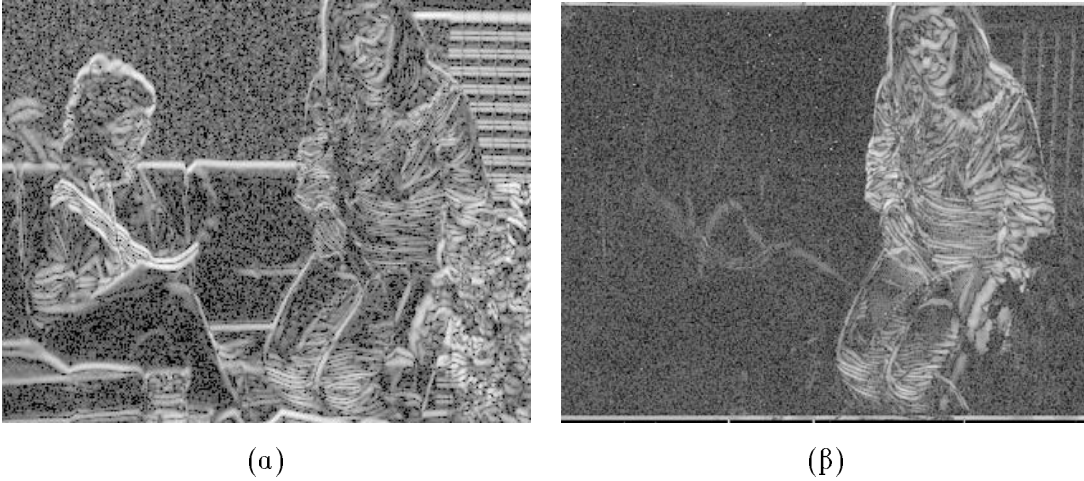
Για την εκτίμηση των παραμέτρων αυτών  $\Theta = \{a_{k,l}; k = 1, 2 \text{ και } l = 1, 2, 3\}$  δεν χρησιμοποιείται μόνο η συνηθισμένη υπόθεση διατήρησης της φωτεινότητας ( $\frac{dI}{dt} = 0$ ) ([37]), αλλά και συνολικές μεταβολές στην φωτεινότητα λαμβάνονται υπόψη με κύριο στόχο την αντιμετώπιση και μεγάλων μετακινήσεων. Έτσι η βασική περιοριστική σχέση μέσω της οποίας εξάγεται η λύση είναι:

$$r(i, j) = I(i + u(i, j), j + v(i, j); t + 1) - I(i, j; t) \approx 0 \quad (4.3)$$

Η εκτίμηση πραγματοποιείται με ένα αυξητικό τρόπο σε ζεύγη εικόνων που έχουν προκύψει με την χρήση πυραμίδας Gauss. Σε κάθε επίπεδο επεξεργασίας εξασφαλίζεται η σύγκλιση με χρήση αλγορίθμων ελαχιστοποίησης, και η λύση αυτή διαχέεται στο επόμενο επίπεδο όπου πάλι εφαρμόζονται αλγόριθμοι σύγκλισης μέχρι το τελευταίο επίπεδο. Εστω  $\hat{\Theta}_k$  η τρέχουσα εκτίμηση για τις παραμέτρους της ίδιας κίνησης στο επίπεδο  $k$ . Η εκτίμηση των παραμέτρων αυτών στο επόμενο επίπεδο ( $k + 1$ ) βασίζεται στην σχέση

$$\hat{\Theta}_{k+1} = \hat{\Theta}'_k + \Delta\Theta_{k+1} \quad (4.4)$$

όπου  $\hat{\Theta}'_k$  οι τιμές των παραμέτρων κίνησης προσαρμοζόμενες στο επίπεδο ( $k + 1$ ) επεξεργασίας. Σε κάθε επίπεδο το πρόβλημα της εκτίμησης των παραμέτρων αυτών ανάγεται στην εκτίμηση του διανύσματος που συσχετίζει με τον καλύτερο δυνατό τρόπο το μοντέλο με τις διαθέσιμες παρατηρήσεις. Για τον λόγο αυτό χρησιμοποιείται M-εκτιμητής ([38]), βασισμένος στην ελαχιστοποίηση ποσότητας αθροισμάτων λάθους ( $\rho(r)$ ) πάνω σε ένα υποσύνολο των σημείων της εικόνας. Η συνάρτηση  $\rho()$  (συνάρτηση προσαρμογής) είναι μία τεχνική διαφοροποίησης του ποσοστού συμμετοχής των διαφόρων σημείων, ανάλογα με την τιμή του  $r(i, j)$ . Έτσι στην



Σχήμα 4.1: Εκτίμηση Ιδίας Κίνησης για την ακολουθία *Interview*

Παράμετροι Ιδίας Κίνησης:  $(\alpha_{1,1}, \alpha_{2,1}) = (5.75, -0.3)$

(α) Εικόνα διαφοράς φωτεινοτήτων.

(β) Εικόνα διαφοράς φωτεινοτήτων μετά την αφαίρεση της Ιδίας Κίνησης.

ελαχιστοποίηση τα σημεία συμμετέχουν με βάρη, γεγονός που έχει σημαντική επίδραση στην εκτίμηση των παραμέτρων του μοντέλου. Η συνάρτηση  $\rho()$  είναι της μορφής,

$$\rho(x) = \begin{cases} x^2, & \text{αν } |x| < 1 \\ 1, & \text{αν } |x| \geq 1 \end{cases} \quad (4.5)$$

Η μορφή της συνάρτησης προσαρμογής είναι αρκετά απλή για λόγους ευκολίας στην υλοποίηση, αλλά υπάρχουν αρκετά πιο πολύπλοκες συναρτήσεις για τον ίδιο λόγο.

Με βάση τα παραπάνω, η λύση του προβλήματος σε κάθε επίπεδο προκύπτει από την ελαχιστοποίηση του παρακάτω αθροίσματος

$$\hat{\Theta} = \min \sum_{(x,y) \in R(t)} T^2 \rho\left(\frac{r(x,y)}{T}\right) \quad (4.6)$$

όπου το  $T$  είναι επίσης άγνωστη μεταβλητή, εξαρτώμενη άμεσα από το είδος και το μέγεθος της κίνησης, αλλά και τις τιμές της παραγώγου της φωτεινότητας στον χρόνο. Η τιμή του  $T$  μπορεί να υπολογιστεί με βάση τις τρέχουσες παραμέτρους της κίνησης,

$$\hat{T} = \frac{\sum_{r(i,j) < T} |r(i,j)|}{\sum_{r(i,j) < T} 1} \quad (4.7)$$

Έτσι στο πρώτο επίπεδο προκύπτει μία αρχική εκτίμηση στην οποία συμμετέχουν όλα τα σημεία (θεωρούμε ότι ο αριθμός των σημείων που αντιστοιχούν στην ίδια κίνηση είναι αρκετά μεγαλύτερος από αυτών που ανήκουν σε ανεξάρτητα κινούμενα αντικείμενα) με μέθοδο ελαχιστών τετραγώνων. Στην συνέχεια και με βάση την συνάρτηση προσαρμογής, στις

εκτιμήσεις συμμετέχουν μόνο τα σημεία που φαίνεται να ανήκουν στην ίδια κίνηση, όπου και πάλι η λύση προκύπτει με μέθοδο ελαχίστων τετραγώνων, με μόνη διαφοροποίηση ότι υπάρχει δυαδικό βάρος για κάθε όρο το οποίο εξαρτάται από την ποσότητα λάθους  $r(i, j)$  και την τιμή του κατωφλιού  $T$ . Ο εκτιμητής αυτός μας επιτρέπει μία καλή εκτίμηση της ίδιας κίνησης, αν αυτή περιγράφεται καλά από το αφηρικό μοντέλο το οποίο έχει επιλεγεί να την αναπαραστήσει στο χώρο προβολής και αν το ποσοστό των σημείων που συμμετέχουν μόνο στην ίδια κίνηση είναι της τάξης του 75%. Στην συνέχεια με την χρήση των παραμέτρων αυτών δημιουργείται μία εξισορροπημένη ακολουθία εικόνων, όπου οι όποιες αλλαγές (μεταβολές) υπάρχουν οφείλονται σε ανεξάρτητα κινούμενα αντικείμενα (Σχήμα 4.1).

### 4.3 Ανίχνευση Ανεξάρτητα Κινούμενων Αντικειμένων

Εστω  $D = \{d(i, j)\}$  η διαφορά των φωτεινιοτήτων μεταξύ των διαδοχικών στιγμιοτύπων στο σημείο  $(i, j)$  με,

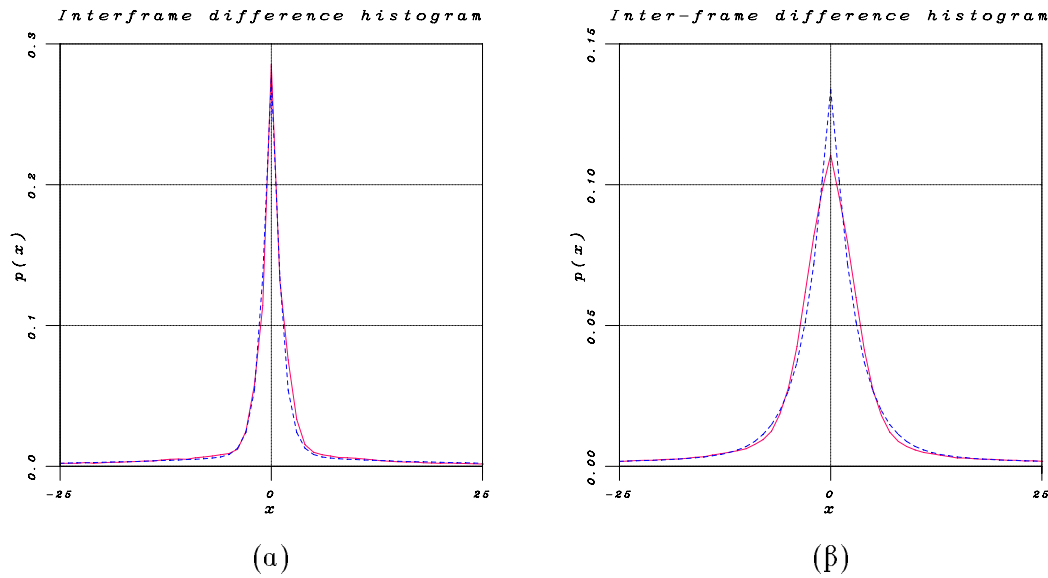
$$d(i, j) = I(i, j; t) - I(i, j; t - 1) \quad (4.8)$$

Όπως έχει ήδη αναφερθεί το πρόβλημα της ανίχνευσης μεταβολών συνίσταται από μία δυαδική απόφαση  $\Theta(i, j)$  για κάθε σημείο στο πλέγμα προβολής στο 2-Δ χώρο. Το τυχαίο πεδίο  $\Theta(i, j)$  αντιστοιχεί σε δύο πιθανά γεγονότα,  $\Theta(i, j) = s$  (Static), αν η παρατηρούμενη διαφορά  $d(i, j)$  υποστηρίζει την υπόθεση για σημείο που ανήκει στο υπόστρωμα ( $H_0$ ), και  $\Theta(i, j) = m$  (Mobile), αν η παρατηρούμενη διαφορά υποστηρίζει την εναλλακτική υπόθεση  $H_1$ , για σημείο που μετακινείται (ανήκει σε ανεξάρτητα κινούμενο αντικείμενο).

Εστω  $p_{D|s}(d|s)$  (αντίστοιχα  $p_{D|m}(d|m)$ ) η συνάρτηση πυκνότητας πιθανότητας της παρατηρούμενης διαφοράς φωτεινιοτήτων κάτω από την υπόθεση  $H_0$  (αντίστοιχα  $H_1$ ). Οι παραπάνω συναρτήσεις πυκνότητας πιθανότητας θεωρούνται ομογενείς, κάτι που σημαίνει ανεξαρτησία ως προς την θέση του σημείου, ενώ συνήθως η μορφή τους είναι Laplace ή Gauss. Στην εργασία αυτή χρησιμοποιούνται κατανομές Laplace (οι οποίες επιβεβαιώνονται από τα δεδομένα) για να περιγράψουν την στατιστική συμπεριφορά των σημείων για τις δύο υποθέσεις, οπότε η δεσμευμένη πυκνότητα πιθανότητας της παρατηρούμενης διαφοράς ως προς κάποια υπόθεση είναι,

$$p_{D|l}(d|l) = \frac{\lambda_l}{2} e^{-\lambda_l |d|}; \quad l \in \{s, m\} \quad (4.9)$$

Η τιμή της παραμέτρου  $\lambda_l$  σχετίζεται με την τυπική απόκλιση  $\sigma_l$  με την σχέση  $\lambda_l = \frac{\sqrt{2}}{\sigma_l}$ . Εστω  $P_s$  (αντίστοιχα  $P_m$ ) η εκ των προτέρων πιθανότητα για την υπόθεση  $H_0$  ( $H_1$ ). Οι τιμές της παρατηρούμενης διαφοράς φωτεινιοτήτων υποθέτουμε ότι προκύπτουν επιλέγοντας μία απόφαση  $l \in \{s, m\}$  με πιθανότητα  $P_l$  και στην συνέχεια επιλέγοντας μία τιμή  $d$  σύμφωνα με την δεσμευμένη συνάρτηση πυκνότητα πιθανότητας  $p_{D|l}(d|l)$ . Έτσι η συνάρτηση πυκνότητα



Σχήμα 4.2: Ανάλυση Ιστογράμματος Διαφοράς Φωτεινότητων με κατανομές Laplace Ακολουθίες (α) *Trevor White*, (β) *Interview* (μετά την αφαίρεση της ίδιας κίνησης)

πιθανότητας για μία δοσμένη τιμή της διαφοράς είναι,

$$p_D(d) = P_s p_{D|s}(d|s) + P_m p_{D|m}(d|m) \quad (4.10)$$

Στην παραπάνω κατανομή (μίξη κατανομών) οι άγνωστες παράμετροι είναι  $\{P_l, \lambda_l; l \in \{s, m\}\}$ . Για την εκτίμηση των παραμέτρων αυτών χρησιμοποιείται η αρχή της Μέγιστης Πιθανοφάνειας ([31] και [52]). Οι άγνωστες παράμετροι εκτιμώνται επαναληπτικά με βάση τις παρατηρούμενες τιμές στη κατανομή της διαφοράς φωτεινότητων. Μία αρχική εκτίμηση (που αποτελεί και την εκκίνηση του αλγορίθμου), προκύπτει με χρήση των ροπών πρώτης, δεύτερης και τρίτης τάξης του παρατηρούμενου δείγματος (Παράρτημα Β). Η ποιότητα της εκτίμησης φαίνεται στο Σχήμα 4.2, όπου υπάρχουν ταυτόχρονα η κατανομή των δειγμάτων αλλά και η προσέγγιση με την μίξη των κατανομών (διακεκομμένη γραμμή) για δύο διαφορετικές ακολουθίες.

Εστω  $A = \{s, m\}$  το σύνολο των πιθανών αποφάσεων, και  $S = \{1, 2, \dots, N\}$  το σύνολο των σημείων της εικόνας. Με βάση τα παραπάνω, ορίζοντας ένα σύστημα γειννίαςσης  $\Gamma$ , το πρόβλημα μπορεί να αντιμετωπιστεί με χρήση μοντέλων Markov, όπου ο χάρτης αποφάσεων περιγράφεται με κατανομές Gibbs. Εστω  $\omega$  ο χάρτης αποφάσεων σε σχέση με το σύνολο  $A$  και  $\omega(i, j) \in A$  η απόφαση που σχετίζεται με το σημείο  $(i, j)$  σύμφωνα με τον χάρτη αποφάσεων  $\omega$ . Στην περίπτωση μας μόνο παρέες δύο σημείων λαμβάνονται υπόψη (Σχήμα 2.1), με κύριο στόχο την μείωση του αριθμού των απαιτούμενων παραμέτρων καθώς και του υπολογιστικού κόστους.

Για να περιγράψουμε την πυκνότητα πιθανότητας  $p(\omega)$ , χρησιμοποιείται κατανομή Gibbs ( $p(\omega) = \frac{1}{Z} e^{-\frac{1}{T} U(\omega)}$ ) ([34]). Η συνάρτησης κόστους  $U(\omega, \tilde{\omega}, d)$  αναλύεται σε άθροισμα τριών

όρων,

$$U(\omega, \tilde{\omega}, d) = U_1(\omega) + U_2(d, \omega) + U_3(\omega, \tilde{\omega}) \quad (4.11)$$

όπου ο όρος  $\tilde{\omega}$  συμβολίζει τον χάρτη αποφάσεων για το προηγούμενο ζεύγος εικόνων της ακολουθίας και

- Ο όρος  $U_1$  εκφράζει την υπόθεση για ομογενή ανακατασκευή, δηλαδή γειτονικά σημεία πρέπει να έχουν παρόμοιες αποφάσεις και είναι της μορφής:

$$U_1(\omega) = \sum_{c \in C_s} V_c(\omega) \quad (4.12)$$

όπου  $C_s$  είναι το σύνολο των δυνατών παρεών (δύο σημείων) στην γειτονιά του  $s$  (με βάση το σύστημα γειτνίασης), ενώ η συνάρτηση δυναμικού  $V_c(\omega)$  δίνεται από τον τύπο:

$$V_c(\omega(i, j), \omega(i+k, j+l)) = \begin{cases} -\alpha_s & , \text{αν } \omega(i, j) = \omega(i+k, j+l) = s \\ -\alpha_m & , \text{αν } \omega(i, j) = \omega(i+k, j+l) = m \\ \alpha_{diff} & , \text{αν } \omega(i, j) \neq \omega(i+k, j+l) \end{cases} \quad (4.13)$$

όπου  $(i+k, j+l)$  γείτονας του  $(i, j)$  (εδώ,  $0 < k^2 + l^2 \leq 2$ ). Στην παραπάνω συνάρτηση,  $\alpha_{diff}$  είναι το κόστος που υπάρχει όταν γειτονικά σημεία έχουν διαφορετικές αποφάσεις,  $\alpha_s$  είναι το δυναμικό που ενισχύει την απόφαση *Static* για γειτονικά σημεία, και  $\alpha_m$  είναι το δυναμικό που ενισχύει την απόφαση *Mobile* κάτω από αντίστοιχες συνθήκες ( $0 < \alpha_{diff} < \alpha_s, \alpha_m$ ).

- Ο όρος  $U_2$  εκφράζει την αντιστοιχία (συνέπεια) μεταξύ των διαθέσιμων παρατηρήσεων και των αποφάσεων και για τον λόγο αυτό χρησιμοποιείται η συνάρτηση πυκνότητας πιθανότητας  $p_{D|\omega(i,j)}(d(i, j)|\omega(i, j))$ , και πιο συγκεκριμένα ισχύει:

$$U_2(d, \omega) = \sum_{(i,j)} -\ln [p_{D|\omega(i,j)}(d(i, j)|\omega(i, j))] \quad (4.14)$$

- Τέλος ο όρος  $U_3$  εκφράζει την συνέπεια με τον χάρτη αποφάσεων των προηγούμενων στιγμιотύπων, και έχει συντηρητικό ρόλο αφού η συνεισφορά του είναι μόνο θετική προς την ελαχιστοποίηση,

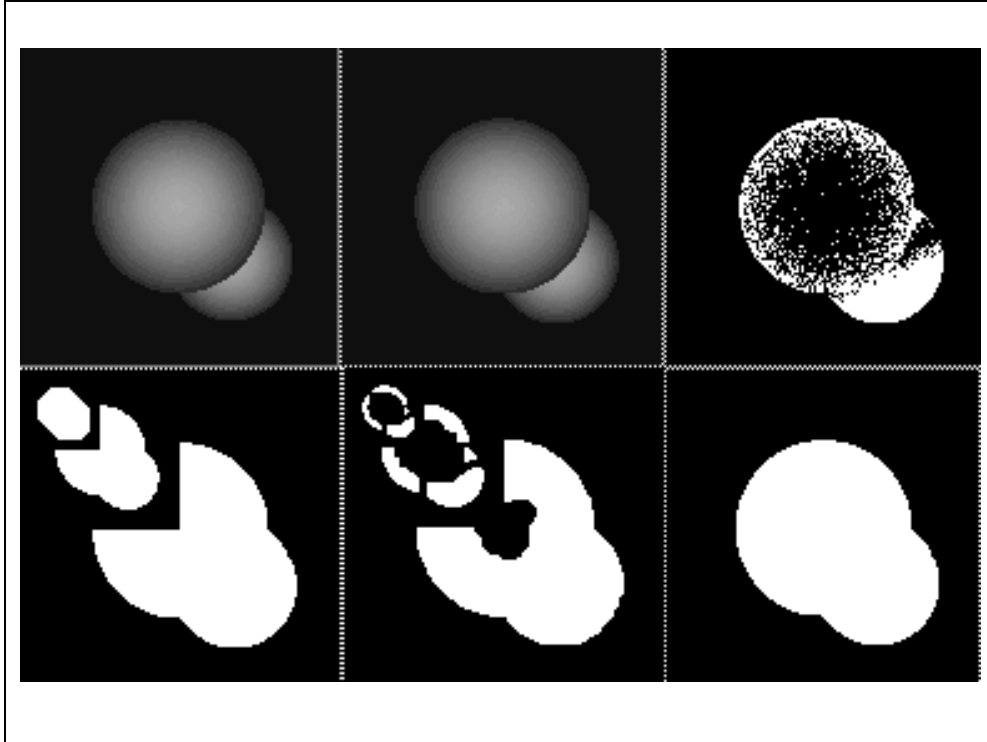
$$U_3(\omega, \tilde{\omega}) = \sum_{s \in S} \rho(\omega(i, j), \tilde{\omega}(i, j)) \quad (4.15)$$

όπου

$$\rho(\omega(i, j), \tilde{\omega}(i, j)) = \begin{cases} -\zeta, & \text{αν } \omega(i, j) = \tilde{\omega}(i, j) \\ 0, & \text{αν } \omega(i, j) \neq \tilde{\omega}(i, j) \end{cases} \quad (4.16)$$

Η λύση του προβλήματος εξάγεται χρησιμοποιώντας κριτήριο μεγιστοποίησης της εκ των υστέρων πιθανότητας (ΜΥΠ), δηλαδή η εκ των υστέρων πιθανότητα ενός χάρτη αποφάσεων με





Σχήμα 4.3: Ανίχνευση Κινούμενων Αντικειμένων: Ακολουθία *Sphera* (Συνθετική)

Πάνω : Εικόνες Εισόδου, Αρχική Εκκίνηση

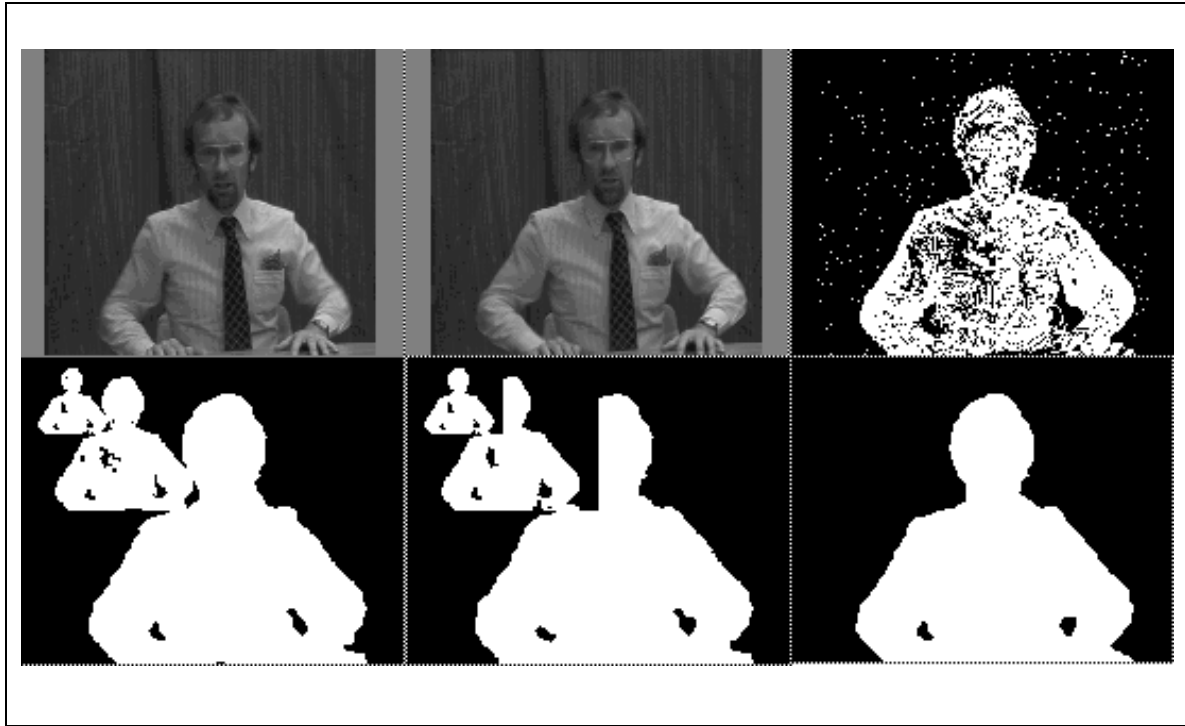
Κάτω : ΕΤΕΠ & ΠΕ (Δεδομένα & Αποφάσεις), ΜΕκΠ & ΠΕ (Αποφάσεις), Βέλτιστο Αποτέλεσμα

δοσμένες τις παρατηρήσεις μεγιστοποιείται, γεγονός που είναι ισοδύναμο με την ελαχιστοποίηση της συνάρτησης κόστους,

$$U(\omega, \tilde{\omega}, d) = \sum_{c \in C_s} V_c(\omega) + \sum_{(i,j)} (\rho(\omega(i,j), \tilde{\omega}(i,j)) - \ln [p_{D|\omega(i,j)}(d(i,j)|\omega(i,j))]) \quad (4.17)$$

Για την ελαχιστοποίηση της παραπάνω ποσότητας δύο διαφορετικοί αυτοκρατικοί αλγόριθμοι καλάρωσης χρησιμοποιούνται σε πολλαπλά επίπεδα με στόχο την μείωση του υπολογιστικού κόστους. Επίσης έχει δοκιμαστεί η αξιοπιστία των αλγορίθμων και σε προσέγγιση ενός επιπέδου. Οι αλγόριθμοι ελαχιστοποίησης ΕΤΕΤ ([11]) και ΜΕκΠ ([20]) σε συνδυασμό με τις δύο τεχνικές πολλαπλών επιπέδων μας δίνουν αρκετά ικανοποιητικές λύσεις στο πρόβλημα.

Ο αλγόριθμος ΕΤΕΤ χρησιμοποιείται ακριβώς με την μορφή που έχει περιγραφεί σε προηγούμενη ενότητα, ενώ για τον αλγόριθμο ΜΕκΠ υπάρχει μία μικρή διαφοροποίηση σε σχέση με την περιγραφή του. Έτσι στις περιπτώσεις που η προσέγγιση είναι σε ένα μόνο επίπεδο (πλήρης ανάλυση των δεδομένων), γίνονται δύο διαδοχικές σαρώσεις αντί για μία, όπου συνήθως στην πρώτη τα δυναμικά που ενισχύουν τις δύο αποφάσεις έχουν την ίδια τιμή ( $\alpha_s = \alpha_m$ ), ενώ στη δεύτερη σάρωση ενισχύεται η απόφαση για κινούμενα σημεία ( $\alpha_s <$



Σχήμα 4.4: Ανίχνευση Κινομένων Αντικειμένων: Ακολουθία *Trevor White*

Πάνω : Εικόνες Εισόδου, Αρχική Εκκίνηση

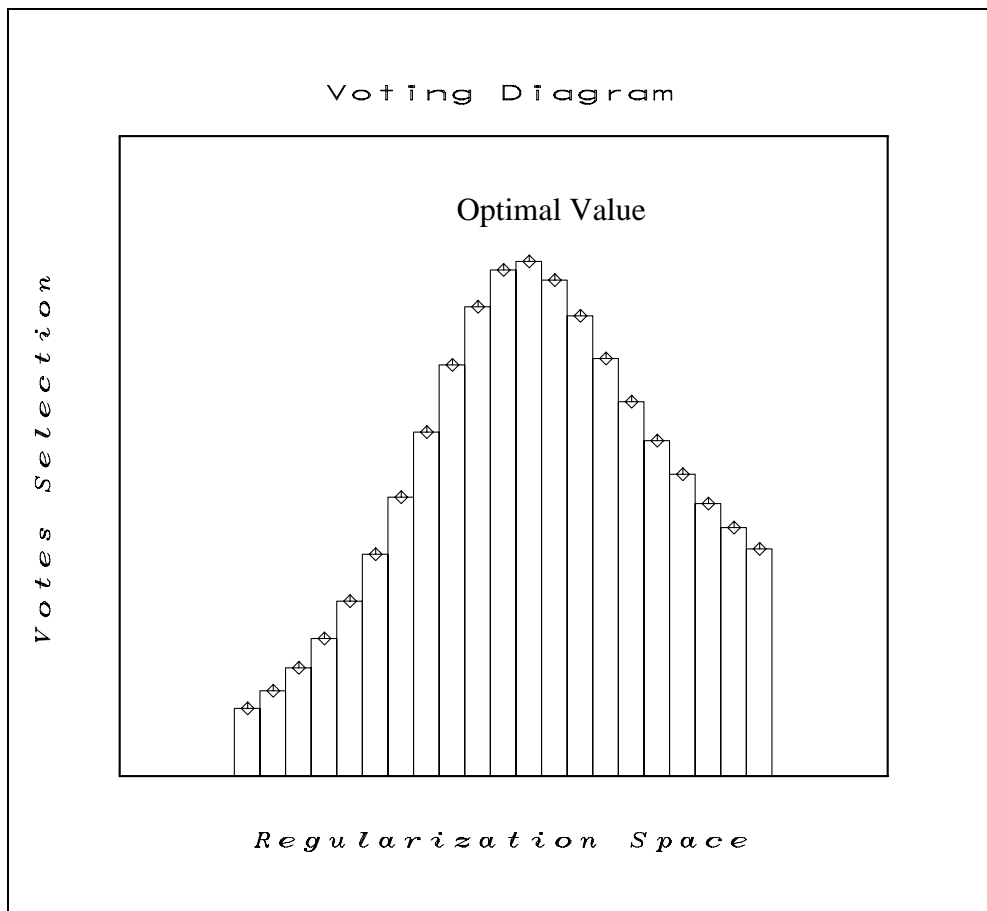
Κάτω : ΕΤΕΠ & ΠΕ (Δεδομένα & Αποφάσεις), ΜΕΚΠ & ΠΕ (Αποφάσεις), Αποτέλεσμα

$\alpha_m$ ). Η εκκίνηση του αλγορίθμου στο δεύτερο πέρασμα είναι ο χάρτης αποφάσεων που έχει προκύψει από το πρώτο πέρασμα. Από πειραματικά αποτελέσματα εξάγεται το συμπέρασμα, ότι η διαδικασία αυτή αποφέρει αρκετά καλά αποτελέσματα και αποτρέπει το φαινόμενο της υπερβολικής λείανσης του χάρτη αποφάσεων.

Από τον ορισμό του προβλήματος είναι εύκολο να διαπιστωθεί ότι αυτό ανάγεται σε μορφή ανακατασκευής με χρήση τεχνικής εξομάλυνσης. Η μόνη διαφορά είναι ότι το μοντέλο είναι πιο πολύπλοκο, και δεν εμφανίζεται μόνο μία παράμετρος στην συνάρτηση δυναμικού. Με στόχο τον έλεγχο της προτεινόμενης μεθόδου για αυτόματο καθορισμό της παραμέτρου εξομάλυνσης, σημαντικές απλοποιήσεις απαιτούνται για την συνάρτηση δυναμικού  $V_c()$ . Έτσι με δεδομένο ότι χρειαζόμαστε μοντέλο μίας παραμέτρου, η ακόλουθη υπόθεση υιοθετείται:

$$\lambda = \alpha_s = \alpha_m = \alpha_{diff} \quad (4.18)$$

Θεωρητικά τουλάχιστον θα περίμενε κανείς ότι αυτή η απλοποίηση θα είχε σημαντική επίδραση στην ευελιξία του μοντέλου, γεγονός που θα έχει και αρνητική επίδραση στα εξαγόμενα αποτελέσματα. Μετά από αυτή την παραδοχή (σχέση (4.18)), το μοντέλο ικανοποιεί τις προϋποθέσεις για της χρήση της προτεινόμενης τεχνικής ψηφοδοσίας για την εκτίμηση της παραμέτρου εξομάλυνσης, όπως αυτή περιγράφεται σε προηγούμενο κεφάλαιο. Για να γίνει



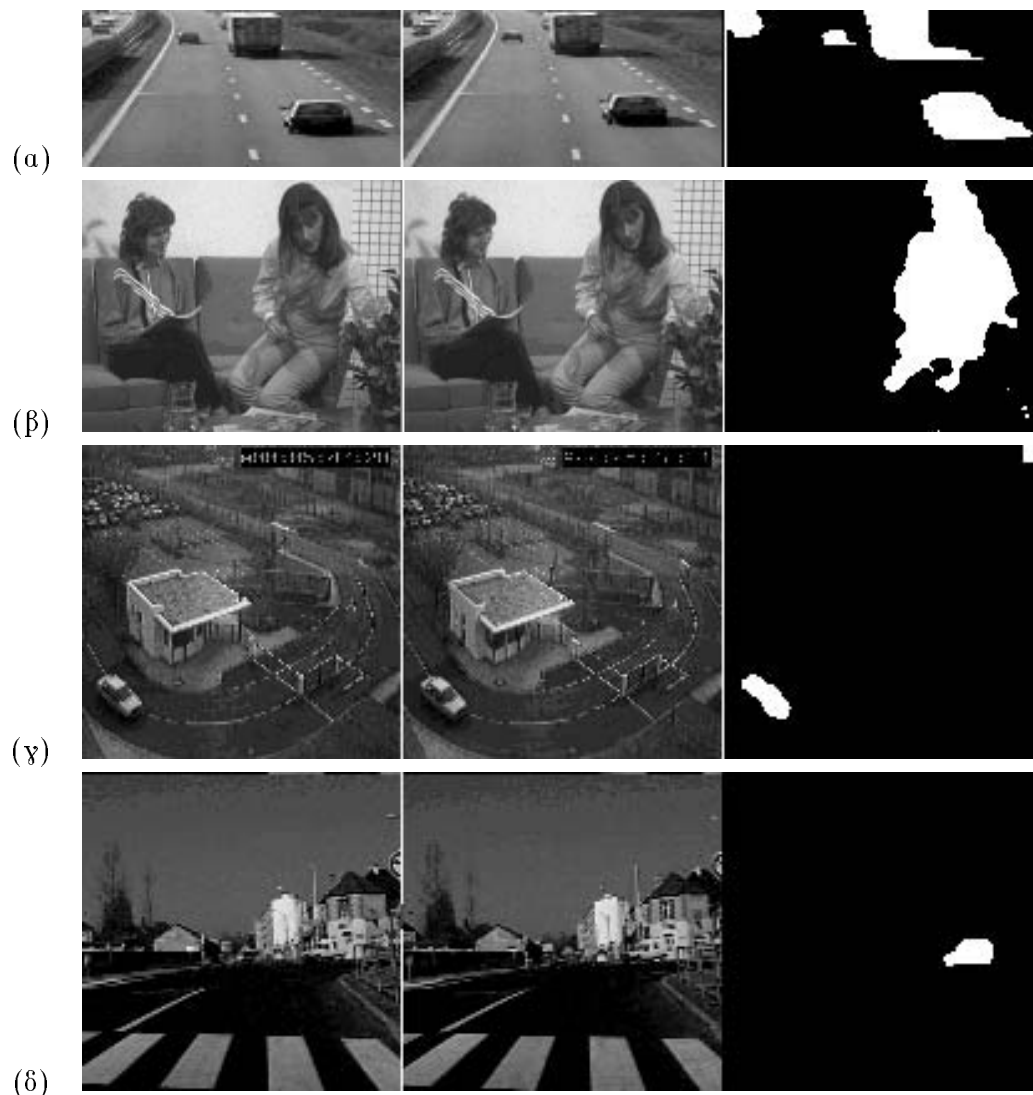
Σχήμα 4.5: Διάγραμμα ψηφοδοσίας για την επιλογή του  $\lambda$

πιο κατανοητή η προτεινόμενη τεχνική, το Σχήμα 4.5 περιέχει ιστόγραμμα ψήφων που αφορά την προτεινόμενη μέθοδο αυτόματου καθορισμού της παράμετρου εξομάλυνσης για την ακολουθία *Trevor White*, ενώ ακολουθεί και σύντομη περιγραφή της διαδικασίας.

Η μέθοδος δουλεύει ως εξής: Αρχικά ο χώρος αναζήτησης του  $\lambda$  θεωρείται το διάστημα  $(0, 2.0)$  με 21 κβαντισμένες τιμές. Ένας αρχικός χάρτης αποφάσεων εξάγεται στο αρχικό επίπεδο της πυραμίδας χρησιμοποιώντας κριτήριο Μέγιστης Πιθανοφάνειας, και η προτεινόμενη μέθοδος εφαρμόζεται. Βέλτιστη τιμή για το  $\lambda$  προκύπτει το 1.1. Στη συνέχεια ο χώρος αναζήτησης υποδιαιρείται και κέντρο του γίνεται αυτή η τιμή (1.1), οπότε ο νέος χώρος αναζήτησης είναι το διάστημα  $(0.6, 1.6)$ . Η τιμή 1.1 χρησιμοποιείται και ένας νέος χάρτης αποφάσεων προκύπτει. Η μέθοδος για τον καθορισμό του  $\lambda$  εφαρμόζεται ξανά και το αποτέλεσμα της μαζί με τον τελευταίο χάρτη αποφάσεων χρησιμοποιούνται σαν αρχικές τιμές στο επόμενο επίπεδο.

Όσον αφορά τα εξαγόμενα πειραματικά αποτελέσματα, είναι αρκετά ικανοποιητικά (Σχήμα 4.6). Έτσι για τις ακολουθίες *Highway*, *Interview*, *Kollning* και *Van* τα αποτελέσματα είναι βέλτιστα. Εδώ θα πρέπει να σημειώσουμε εδώ ότι τα αποτελέσματα αυτά προέκυψαν αυτόματα,

χωρίς την συμμετοχή του χρήστη σε κάποιο στάδιο της επεξεργασίας. Επιπλέον σχολιασμός της προτεινόμενης μεθόδου, και πιο εμπειροταωμένη ανάλυση των αποτελεσμάτων που μας δίνει, υπάρχουν στο Κεφάλαιο 6. Τέλος για λόγους σύγκρισης παρατίθενται και τα αποτελέσματα που προέκυψαν με τιμή 1.0 για την παράμετρο εξομάλυνσης στις ίδιες ακολουθίες (Σχήμα 4.7)



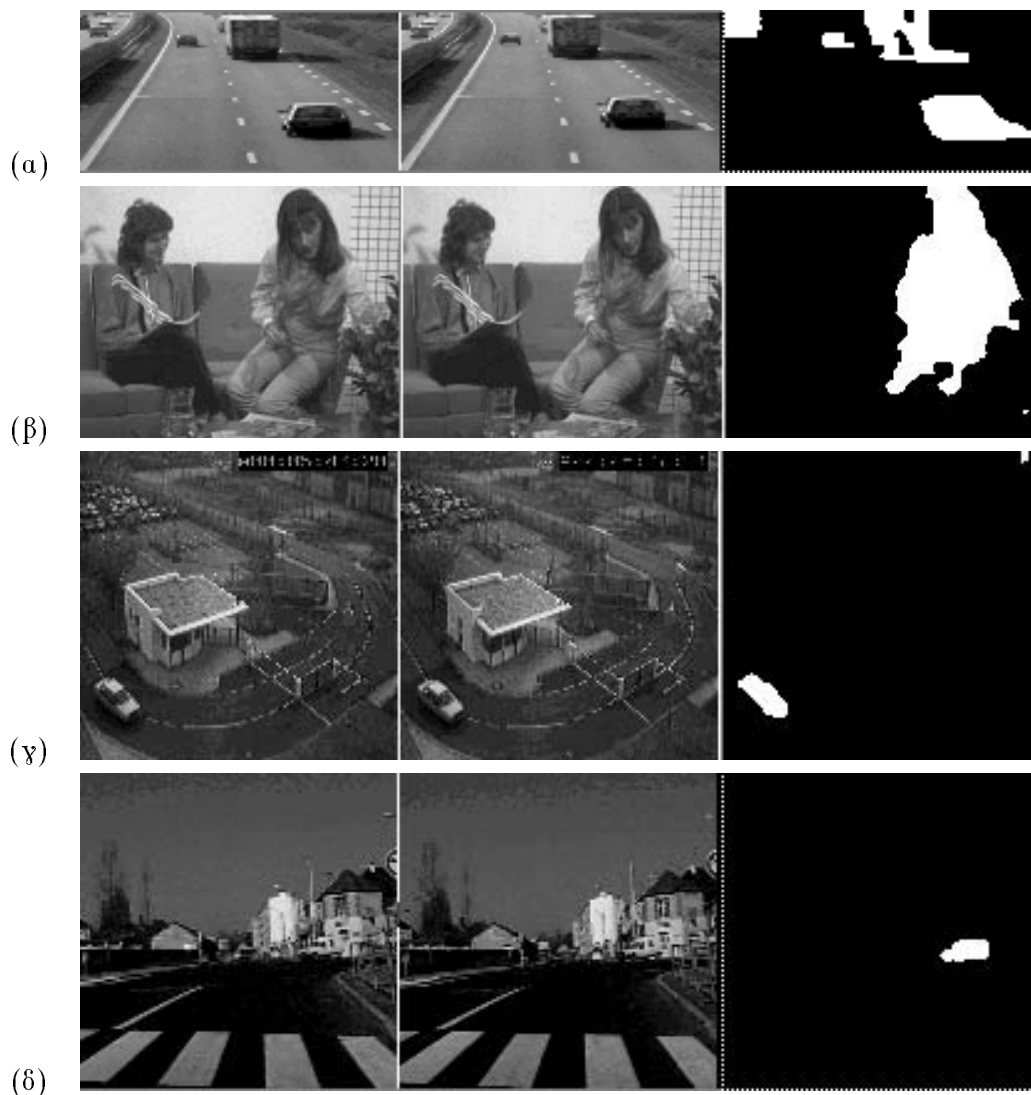
Σχήμα 4.6: Δυναμική Ανίχνευση Κινούμενων Αντικειμένων

(α) Ακολουθία *Highway* (ΕΤΕΤ),  $\lambda=1.0125$

(β) Ακολουθία *Interview* - Κινούμενος Παρατηρητής (ΜΕκΠ),  $\lambda=0.8125$

(γ) Ακολουθία *Kollnig* (ΜΕκΠ),  $\lambda=0.8625$

(δ) Ακολουθία *Van* (ΕΤΕΤ),  $\lambda=0.9875$



Σχήμα 4.7: Ανίχνευση Κινομένων Αντικειμένων,  $\lambda=1.00$

- (α) Ακολουθία *Highway* (ΕΤΕΤ)
- (β) Ακολουθία *Interview* - Κινοόμενος Παρατηρητής (ΜΕκΠ)
- (γ) Ακολουθία *Kollnig* (ΜΕκΠ)
- (δ) Ακολουθία *Van* (ΕΤΕΤ)

## Κεφάλαιο 5

# Εντοπισμός Θέσης Κινούμενων Αντικειμένων

### 5.1 Εισαγωγή-Ανασκόπηση

Η μοντελοποίηση του προβλήματος εντοπισμού της θέσης των αντικειμένων είναι παρόμοια με αυτή που ακολουθήθηκε στην ανίχνευση μεταβολών. Η ανακατασκευή σε αυτή την περίπτωση είναι πιο πολύπλοκη, αφού ο στόχος είναι ο χαρακτηρισμός της κατάστασης που είναι πιο πιθανή και στις δύο εικόνες για κάθε σημείο  $(i, j)$ . Κάθε σημείο σε κάθε εικόνα μπορεί να είναι είτε σημείο του υποστρώματος είτε σημείο που ανήκει σε κινούμενο αντικείμενο. Εστω  $L = \{B, O\}$  το σύνολο των δυνατών αποφάσεων που αφορούν την κατάσταση του σημείου, όπου το  $B$  συμβολίζει το υπόστρωμα (background) ενώ το  $O$  αντιστοιχεί σε αντικείμενο (object). Στο πρόβλημα του εντοπισμού της θέσης των αντικειμένων, ένα ζεύγος αποφάσεων πρέπει να εκτιμηθεί για κάθε σημείο  $(\Theta(i, j; t), \Theta(i, j; t+1)) \in L \times L$ . Ο συμβολισμός αυτός είναι ισοδύναμος με την απόφαση  $\Theta(i, j; t)$  (αντίστοιχα  $\Theta(i, j; t+1)$ ) να περιγράφει την κατάσταση που ισχύει για το σημείο  $(i, j)$  στην εικόνα της χρονικής στιγμής  $t$  (αντίστοιχα  $t+1$ ). Με βάση τα παραπάνω έχουμε τέσσερις δυνατές καταστάσεις και πιο συγκεκριμένα ισχύει:

$$\begin{aligned} H_{00} &: (\Theta(i, j; t), \Theta(i, j; t+1)) = (B, B) \\ H_{01} &: (\Theta(i, j; t), \Theta(i, j; t+1)) = (B, O) \\ H_{10} &: (\Theta(i, j; t), \Theta(i, j; t+1)) = (O, B) \\ H_{11} &: (\Theta(i, j; t), \Theta(i, j; t+1)) = (O, O) \end{aligned} \tag{5.1}$$

Το παραπάνω πρόβλημα δεν έχει διερευνηθεί στο βαθμό που ισχύει για το πρόβλημα της ανίχνευσης μεταβολών. Αν και η αντιμετώπιση του βρίσκει αρκετές εφαρμογές σε πολλές διεργασίες της Μηχανικής Ορασης και Μάθησης, το απαγορευτικό κόστος καθώς και η πολυπλοκότητά του όσον αφορά την μοντελοποίηση του, είναι οι κύριες αιτίες για τον μικρό

αριθμό προσεγγίσεων που υπάρχουν στην βιβλιογραφία. Στην κατάσταση αυτή συντελεί και το γεγονός ότι η δειγματοληψία των εικόνων έχει εξελιχθεί σε τέτοιο βαθμό, ώστε η λύση τείνει να είναι ισοδύναμη με αυτή της ανίχνευσης μεταβολών.

Το πρόβλημα αρχικά αντιμετωπίστηκε από τους P. Bouthemy και P. Lalande ([16] και [17]). Η προτεινόμενη μέθοδος βασίζεται στην χρήση τριών διαδοχικών στιγμιότυπων της ακολουθίας, ενώ η λύση προκύπτει με την ελαχιστοποίηση συνάρτησης κόστους με χρήση επαναληπτικού αιτιοκρατικού οχήματος χαλάρωσης. Η ιδέα στην οποία βασίζεται η προσέγγιση είναι χρήση τριών εικόνων, στις χρονικές στιγμές  $t_1$ ,  $t_2$ ,  $t_3$  για τον εντοπισμό των κινούμενων αντικειμένων στην χρονική στιγμή  $t_2$ . Δύο δυαδικοί χάρτες αποφάσεων που αφορούν το πρόβλημα της ανίχνευσης μεταβολών για την κίνηση μεταξύ των στιγμιότυπων  $t_1$  και  $t_2$ , και,  $t_2$  και  $t_3$  υπολογίζονται και στην συνέχεια η βασική δυαδική πράξη “ΚΑΙ” (AND) μεταξύ τους μας δίνει τις θέσεις των αντικειμένων στο στιγμιότυπο  $t_2$ . Τα κυριότερα μειονεκτήματα της μεθόδου, είναι ο όγκος των διαθέσιμων πληροφοριών που απαιτείται για την λύση του προβλήματος, αφού για να προκύψει η θέση των αντικειμένων σε κάποια εικόνα απαιτούνται τρία διαδοχικά στιγμιότυπα της ακολουθίας, καθώς επίσης και η λύση του προβλήματος ανίχνευσης μεταβολών σε δύο διαφορετικά ζεύγη εικόνων.

Αρκετά ενδιαφέρουσα προσέγγιση του προβλήματος, παρόμοια με αυτή που θα ακολουθηθεί και στην προτεινόμενη λύση έχει επίσης προταθεί ([40]). Το πρόβλημα μοντελοποιείται με χρήση κρυφών πεδίων Markov για την ανίχνευση της κίνησης, όπου χρησιμοποιούνται δύο ειδών παρατηρήσεις. Οι διαφορές στην φωτεινότητα καθώς και μία τιμή που προκύπτει με την χρήση μάσκας τοπικών αλλαγών, δηλαδή ο διαχωρισμός των σημείων που ανήκουν σε κινούμενα αντικείμενα και αυτών που ανήκουν στο υπόστρωμα. Μία επιπλέον πληροφορία που χρησιμοποιείται για την βελτίωση της απόδοσης του αλγορίθμου είναι η φωτεινότητα στο δοσμένο σημείο. Ετσι τέσσερις διοδιάστατες κατανομές Gauss χρησιμοποιούνται για να περιγράψουν την αντιστοιχία μεταξύ αποφάσεων και δεδομένων. Το κυριότερο μειονέκτημα της μεθόδου είναι ο μεγάλος αριθμός των παραμέτρων που απαιτούνται για την περιγραφή των παραπάνω κατανομών (25 παράμετροι), καθώς επίσης και η υπόθεση ότι οι φωτεινότητες στα κινούμενα αντικείμενα είναι ομοιόμορφες (ώστε να είναι δυνατή η μοντελοποίηση τους άμεσα με την χρήση μίας και μόνο κατανομής Gauss). Η πολυπλοκότητα του μοντέλου αυξάνει αρκετά αν υπολογίσουμε και τις παραμέτρους που απαιτούνται για το μοντέλο Markov, με αποτέλεσμα να φτάνουμε στον αρκετά μεγάλο αριθμό των 32 παραμέτρων (μερικές από αυτές είναι οι ίδιες). Τέλος αρκετά επισφαλής είναι και ο τρόπος καθορισμού μερικών παραμέτρων αυτών με χρήση “ακολουθίας διδασκαλίας” όπου σε ένα μικρό τμήμα της ακολουθίας καθορίζονται εμπειρικά οι μάσκες που μας δίνουν τα σημεία που αντιστοιχούν σε κίνηση. Δηλαδή ο χρήστης έχει άμεση συμμετοχή στην φάση της επεξεργασίας. Στην συνέχεια οι τιμές των παραμέτρων των κατανομών Gauss καθορίζονται με την χρήση κριτηρίου Μέγιστης



Πιθανοφάνειας. Τέλος αξιοσημείωτο είναι το γεγονός ότι η προσέγγιση αντιμετωπίζει το πρόβλημα με δεδομένα που προέρχονται από δύο μόνο διαδοχικά στιγμιότυπα της ακολουθίας.

Η στατιστική οριοθέτηση του προβλήματος στην εργασία αυτή γίνεται με χρήση πεδίων Markov όπου κατανομές Gibbs περιγράφουν το πρόβλημα της τμηματοποίησης. Το κυριότερο πρόβλημα είναι η δημιουργία ενός δυναμικού στατιστικού μοντέλου ικανού να περιγράψει την αντιστοιχία μεταξύ των φωτεινότητων στο ζεύγος των εικόνων και των αποφάσεων που παίρνονται. Για τον λόγο αυτό είτε κατανομές Gauss, είτε τμηματικά ομοιόμορφες κατανομές (με άμεσο στόχο την μείωση του κόστους) προσεγγίζουν τις παρατηρούμενες φωτεινότητες σε συνάρτηση με την ανίχνευση μεταβολών. Στην συνέχεια με βάση την παραπάνω ανάλυση δημιουργείται ένα δυναμικό μοντέλο ικανό να μας δώσει ένα μέτρο ποιότητας των αποφάσεων που λαμβάνονται για το ζεύγος των εικόνων σε δοσμένο σημείο, σε σχέση με τις διαθέσιμες παρατηρήσεις στο σημείο αυτό. Συναρτήσεις κόστους κατασκευάζονται βασιζόμενες στις παραπάνω παρατηρήσεις και στην υπόθεση για ομοιογενή ανακατασκευή. Η λύση του προβλήματος προκύπτει από ελαχιστοποίηση των συναρτήσεων αυτών.

## 5.2 Στατιστική Ανάλυση

Ας επανέλθουμε τώρα στο ορισμό και την μοντελοποίηση του προβλήματος. Το σύνολο των διαθέσιμων παρατηρήσεων αποτελείται από την τμηματοποίηση όσον αφορά την ανίχνευση μεταβολών, καθώς και τις φωτεινότητες στις δύο εικόνες. Ο πρώτος στόχος είναι η δημιουργία ενός όρου που εκφράζει την αντιστοιχία μεταξύ των δεδομένων και των αποφάσεων και για τον λόγο αυτό θα χρησιμοποιηθούν δεομευμένες συνάρτησεις πυκνότητας πιθανότητας. Έστω

$$p((I(i, j; t), I(i, j; t + 1)) = (x_0, x_1) | (\Theta(i, j; t), \Theta(i, j; t + 1)) = (\alpha, \beta) \quad (5.2)$$

η δεομευμένη συνάρτηση πυκνότητας πιθανότητας για την περίπτωση  $(\alpha, \beta)$ , όπου  $(\alpha, \beta) \in L \times L$  και  $I(i, j; t)$  η τιμή της φωτεινότητας στο σημείο  $(i, j)$ . Στην περίπτωση που  $\alpha \neq \beta$  η επίλυση του προβλήματος είναι απλή αφού τα δύο γεγονότα είναι ανεξάρτητα μεταξύ τους και η διοδιάστατη συνάρτηση πυκνότητας πιθανότητας μπορεί να προκύψει άμεσα σαν γινόμενο των μονοδιάστατων, δηλαδή

$$p(x_0, x_1 | \alpha, \beta) = p(x_0 | \alpha) p(x_1 | \beta) \quad (5.3)$$

Με βάση την παραπάνω υπόθεση είναι σχετικά εύκολος ο υπολογισμός της πιθανότητας για τις περιπτώσεις  $(B, O), (O, B)$ , με γνωστές τις τιμές των φωτεινότητων στο δοσμένο σημείο, μπορεί να εξαχθεί με άμεσο τρόπο από τις μονοδιάστατες κατανομές. Το πρόβλημα που υπάρχει είναι αφενός ο υπολογισμός των μονοδιάστατων συναρτήσεων πυκνότητας πιθανότητας για τις δύο δυνατές καταστάσεις και αφετέρου ο τρόπος υπολογισμού των πιθανοτήτων στις περιπτώσεις  $(B, B), (O, O)$ .

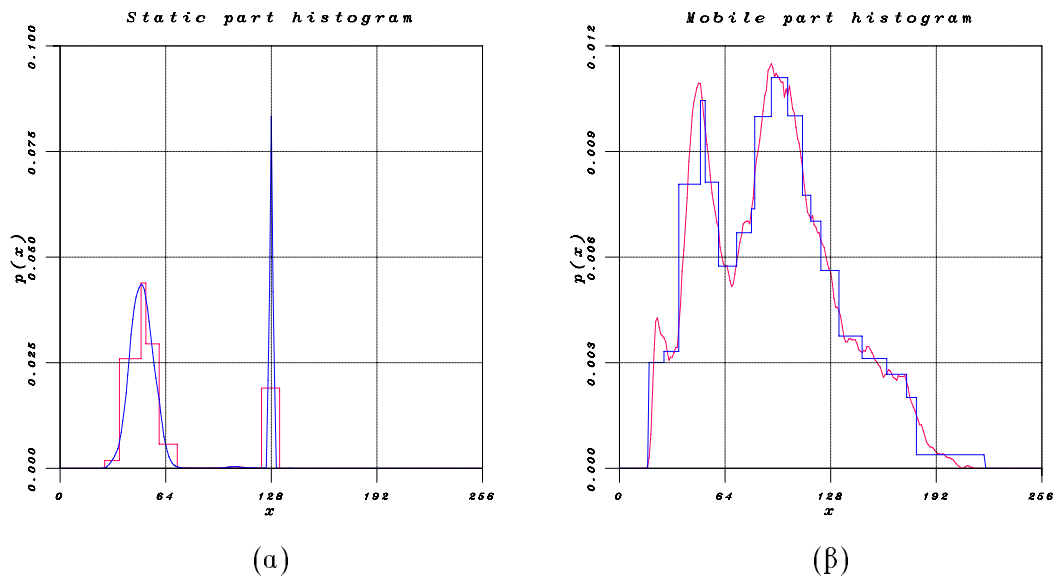
Με βάση την τμηματοποίηση όσον αφορά το πρόβλημα της ανίχνευσης μεταβολών είναι δυνατόν να προκύψουν δύο διαφορετικά ιστογράμματα φωτεινότητας, ένα που περιγράφει τα σημεία που δεν κινούνται (Static), και ένα άλλο για τα υπόλοιπα. Τα ιστογράμματα αυτά είναι πολύ πιθανό να αποτελούνται από διαφορετικούς πληθυσμούς ανάλογα με τις τιμές στην φωτεινότητα (αν υποθέσουμε ότι ένας άνθρωπος κινείται τότε οι τιμές της φωτεινότητας στο χώρο προβολής δεν ακολουθούν ομοιόμορφη κατανομή, αφού είναι δυνατόν να φοράει ρούχα διαφορετικού χρώματος). Χρησιμοποιώντας τον χάρτη αποφάσεων που αφορά το πρόβλημα της ανίχνευσης μεταβολών και τα σημεία που ανήκουν σε περιοχές που δεν υπάρχει κίνηση είναι εφικτός ο υπολογισμός του ιστογράμματος που αφορά αυτό το τμήμα της εικόνας. Σημαντική παρατήρηση και διαφοροποίηση του μοντέλου, είναι ότι τα σημεία για τα οποία η διαφορά των φωτεινότητων στις δύο εικόνες είναι αρκετά μεγάλη, εξαιρούνται και δεν χρησιμοποιούνται για την κατασκευή του ιστογράμματος που αντιστοιχεί στην κίνηση. Αυτό γίνεται γιατί τα σημεία αυτά είναι πολύ πιθανό να ανήκουν σε επικαλυπτόμενα ή αποκαλυπτόμενα τμήματα. Η ύπαρξη θορύβου είναι συχνό φαινόμενο σε τέτοια προβλήματα, και για να αποφευχθούν τυχόν παρενέργειες στο μοντέλο εξαιτίας του, μία διαδικασία λείανσης εφαρμόζεται στα παρατηρούμενα ιστογράμματα. Το πρόβλημα ανάγεται στην χρήση αυτών των ιστογραμμάτων για τον υπολογισμό των παραπάνω πιθανοτήτων.

Στην συνέχεια ακολουθούν δύο διαφορετικές λύσεις στο πρόβλημα της εκτίμησης των πιθανοτήτων αυτών, η πρώτη βασισμένη σε τμηματικά ομοιόμορφες κατανομές και η δεύτερη σε τεχνικές ανάλυσης μίξης κατανομών.

### 5.2.1 Πιθανοτική προσέγγιση με τμηματικά ομοιόμορφες κατανομές

Άμεσος τρόπος για την λύση του προβλήματος αυτού είναι η προσέγγιση των δύο τυχαίων μεταβλητών (στατικού και κινούμενου τμήματος) που περιγράφουν τη σχέση αποφάσεων-φωτεινότητας, με τμηματικά ομοιόμορφες κατανομές. Έτσι χρησιμοποιείται ένα ομοιόμορφο μοντέλο με σταθερές τιμές στα διάφορα κβαντισμένα διαστήματα για τις συναρτήσεις πυκνότητας πιθανότητας. Η τεχνική αυτή έχει βασικό πλεονέκτημα το μικρό υπολογιστικό κόστος, και όπως φαίνεται και στην συνέχεια η σχέση κόστους-απόδοσης είναι αρκετά καλή. Ας προχωρήσουμε όμως σε μία πιο λεπτομερή περιγραφή του μοντέλου.

Η γενική ιδέα είναι η διαίρεση του συνόλου των δυνατών τιμών των φωτεινότητων σε μη επικαλυπτόμενα διαστήματα. Η διαδικασία αυτή πρέπει να γίνει με τέτοιο τρόπο ώστε η συνάρτηση πυκνότητας πιθανότητας για κάποια απόφαση με δοσμένη την φωτεινότητα να παρουσιάζει ομοιόμορφες τιμές σε κάθε διάστημα, γεγονός που μας επιτρέπει την χρήση μίας αντιπροσωπευτικής τιμής για την πιθανότητα ως προς κάποια απόφαση στο διάστημα αυτό. Η παραπάνω υπόθεση πρέπει να ισχύει και για τις δύο δυνατές καταστάσεις. Στην συνέχεια οι πιθανότητες για το διοδιάστατο διάνυσμα αποφάσεων μπορούν να προκύψουν από

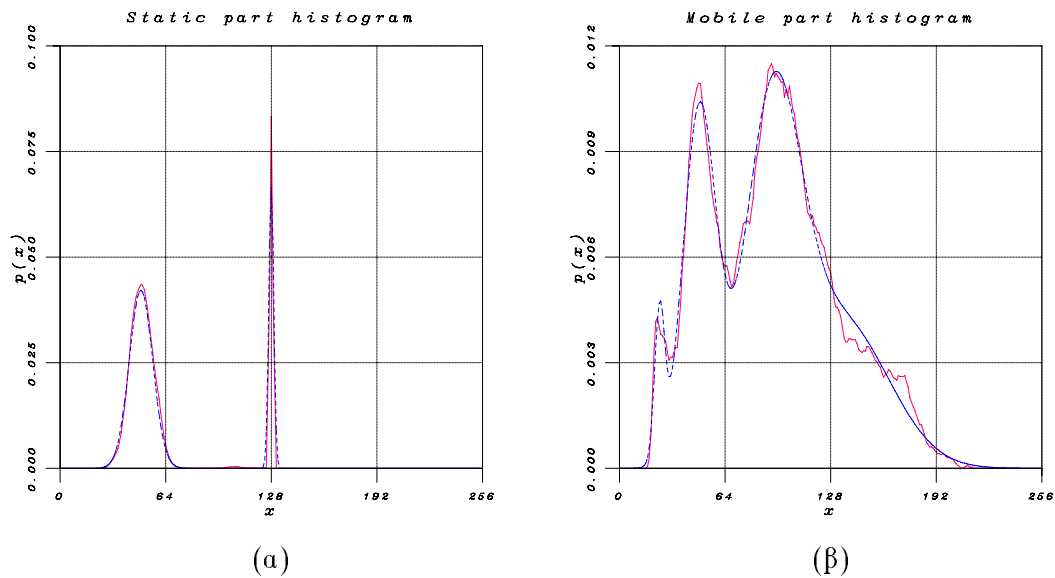
Σχήμα 5.1: Ακολουθία *Trevor White* - Ανάλυση Ιστογράμματος

Τμηματικά Ομοιόμορφες Κατανομές: (α) Στατικό Τμήμα, (β) Κινούμενο Τμήμα

την ορθογώνια διαίρεση του διοδιάστατου χώρου τιμών για το ζεύγος φωτεινοτήτων στις δύο εικόνες. Αφού η διαίρεση του διοδιάστατου χώρου των φωτεινοτήτων γίνεται ορθογώνια για να καλύπτει τις περιπτώσεις των δύο ανεξάρτητων κατανομών  $(B, O)$  η προσέγγιση με τμηματικά ομοιόμορφες κατανομές μπορεί να γίνει στον μονοδιάστατο χώρο και στην συνέχεια είναι εύκολο να προκύψουν οι ζητούμενες πιθανότητες για τα ζεύγη των αποφάσεων.

Η λύση του προβλήματος της ανίχνευσης μεταβολών και η αναγκαιότητα μίας καλής αναπαράστασης τόσο του κινούμενου τμήματος των εικόνων όσο και του στατικού, ανεξάρτητα από το εύρος των τιμών των φωτεινοτήτων, οδηγούν στην προσέγγιση του προβλήματος με δύο διαφορετικούς κβαντιστές, έναν για κάθε υπόθεση. Με την τεχνική αυτή προκύπτουν δύο διαφορετικές προσεγγίσεις με τμηματικά ομοιόμορφες κατανομές, για το ίδιο σύνολο τιμών με κριτήριο την πιθανότητα κάποιας φωτεινότητας δεδομένης κάποιας απόφασης. Ο στόχος είναι ο χωρισμός του συνόλου τιμών σε διαστήματα, με τέτοιο τρόπο ώστε και οι δύο καταστάσεις να έχουν σταθερή συμπεριφορά όσον αφορά την πιθανότητα απόφασης. Δηλαδή να μπορούν να χρησιμοποιηθούν ομοιόμορφες κατανομές για να προσεγγίσουν τις παρατηρούμενες κατανομές στα διαστήματα αυτά. Για τον λόγο αυτό απαιτείται στη συνέχεια ενοποίηση των δύο κβαντιστών σε ένα, του οποίου τα επίπεδα απόφασης προκύπτουν από την ένωση των επιπέδων απόφασης των δύο αρχικών.

Ένα σημαντικό πρόβλημα στην παραπάνω διαδικασία είναι ο καθορισμός των επιπέδων που θα υπάρχουν σε κάθε κβαντιστή. Το πρόβλημα αυτό αντιμετωπίζεται με χρήση των διαθέσιμων πληροφοριών από τα ιστογράμματα καθώς και με την χρήση κριτηρίου ελαχιστοποίησης μέσου



Σχήμα 5.2: Ακολουθία *Trevor White* - Ανάλυση Ιστογράμματος  
Κατανομές Gauss : (α) Στατικό Τμήμα, (β) Κινούμενο Τμήμα

τετραγωνικού σφάλματος στον κβαντισμό. Αρχικά επιλέγεται το πλήθος των επιπέδων με βάση πιθανά μέγιστα στο ιστόγραμμα. Στην συνέχεια εφαρμόζεται ο αλγόριθμος των Lloyd-Max ([35]) μέχρις ότου να εξασφαλιστεί η σύγκλιση. Αν το συνολικό μέσο τετραγωνικό σφάλμα είναι πάνω από κάποιο εμπειρικά καθοριζόμενο κατώφλι, το επίπεδο με το μεγαλύτερο τοπικό μέσο τετραγωνικό σφάλμα υποδιαιρείται και ο αλγόριθμος των Lloyd-Max εφαρμόζεται ξανά με ένα επίπεδο κβαντισμού παραπάνω. Η διαδικασία αυτή επαναλαμβάνεται μέχρι το συνολικό μέσο τετραγωνικό σφάλμα να πάρει τιμή κάτω από το δοσμένο κατώφλι.

Στην συνέχεια με βάση τις τελικές αποφάσεις όσον αφορά τα επίπεδα κβαντισμού και τα δοσμένα ιστογράμματα, υπολογίζεται η πιθανότητα σε κάθε επίπεδο και για τις δύο αποφάσεις (Σχήμα 5.1). Έχοντας τις πιθανότητες αυτές μπορούμε άμεσα να υπολογίσουμε τις πιθανότητες των καταστάσεων  $(B, O)$  και  $(O, B)$  σαν γινόμενο των επιμέρους μονοδιάστατων πιθανοτήτων, ενώ για τις άλλες δύο καταστάσεις οι πιθανότητες προκύπτουν άμεσα από το διδιάστατο ιστόγραμμα των φωτεινοτήτων πάνω στα κβαντισμένα διαστήματα.

### 5.2.2 Πιθανοτική προσέγγιση με κατανομές Gauss

Μία πιο αξιόπιστη προσέγγιση είναι ανάλυση των παρατηρούμενων κατανομών (για τις δύο υποθέσεις) σαν μίξη κατανομών, όπως γίνεται στην ανίχνευση μεταβολών για το ιστόγραμμα της διαφοράς φωτεινοτήτων. Η αξιοπιστία της μεθόδου οφείλεται στο ότι τα παρατηρούμενα ιστογράμματα προσεγγίζονται με μεγαλύτερη ακρίβεια με μίξη κατανομών, σε σχέση με την ακρίβεια που προκύπτει με την χρήση τμηματικά ομοιόμορφων κατανομών. Στην προκειμένη

περίπτωση χρησιμοποιούνται κατανομές Gauss και με βάση αυτή την υπόθεση, η συνάρτηση πυκνότητας πιθανότητας για τις τιμές των φωτεινοτήτων τόσο για το στατικό τμήμα της εικόνας όσο και για το κινούμενο, μπορεί να αναλυθεί σαν μίξη τέτοιων κατανομών δηλαδή,

$$p(x|\alpha) = \sum_{i=1}^{c_\alpha} \frac{P_{\alpha i}}{\sigma_{\alpha i} \sqrt{2\pi}} e^{-\frac{(x-\mu_{\alpha i})^2}{2\sigma_{\alpha i}^2}} \quad (5.4)$$

Το πρόβλημα στην συνέχεια ανάγεται στην εκτίμηση των παραμέτρων που συνθέτουν το παραπάνω δείγμα. Αρχική προϋπόθεση για την λύση του προβλήματος είναι ο καθορισμός του αριθμού των πληθυσμών  $c_\alpha$  που συνθέτουν το παραπάνω δείγμα, παράμετρος που επιδρά σημαντικά στη συνέχεια στην εξαγωγή των παραμέτρων των κατανομών. Ο αριθμός αυτός εξάγεται εμπειρικά με βάση το παρατηρούμενο ιστόγραμμα. Έτσι στο ιστόγραμμα αυτό αναζητούμε τοπικά μέγιστα τα οποία μπορούμε να υποθέσουμε ότι αρχικά αντιστοιχούν και σε μέσες τιμές διαφορετικών πληθυσμών. Η προσέλαση των τοπικών μεγίστων γίνεται ιεραρχικά και επιπλέον απαιτούμε να υπάρχει ελάχιστο κάτω φράγμα στην απόσταση μεταξύ κάθε υποψήφιας τιμής και των ήδη επιλεγθέντων μεγίστων. Έτσι αποφεύγεται η διαίρεση ίδιων πληθυσμών που παρουσιάζουν μεγάλη πιθανότητα. Στην συνέχεια με την χρήση κριτηρίου Μέγιστης Πιθανοφάνειας για την εκτίμηση των παραμέτρων αυτών, μπορούμε να εκτιμήσουμε τις άγνωστες παραμέτρους  $(P_{Bi}, \sigma_{Bi}, \mu_{Bi})$  για κάθε πληθυσμό. Η ίδια ακριβώς προσέγγιση χρησιμοποιείται για τον υπολογισμό του  $p(x|O)$ . Η παραπάνω διαδικασία καθώς και η κατασκευή των ιστογραμμάτων γίνεται μόνο στη μία από τις δύο εικόνες, αφού υποθέτουμε ότι δεν υπάρχει σημαντική διαφορά μεταξύ των δύο εικόνων, όσον αφορά το περιεχόμενο. Επομένως και οι αντίστοιχες μεταβλητές είναι σταθερές στον χρόνο. Τα αποτελέσματα αυτής της προσέγγισης είναι αρκετά ικανοποιητικά και η απόκλιση μεταξύ προσέγγισης και δείγματος είναι πολύ μικρή (Σχήμα 5.2).

Το πρόβλημα παραμένει για τις περιπτώσεις  $(B, B), (O, O)$  και δύο διαφορετικές λύσεις προτείνονται. Αρκετά απλή και με σχετικά μικρό υπολογιστικό κόστος είναι ο υπολογισμός ενός συνολικού συντελεστή συσχέτισης  $\rho_\alpha$ , για κάθε περίπτωση. Στην συνέχεια με την χρήση των παραπάνω συντελεστών και υποθέτοντας ότι αυτοί ισχύουν ανεξάρτητα από τους πληθυσμούς που συνθέτουν την κατανομή των φωτεινοτήτων, μπορούμε να γράψουμε

$$p(x_0, x_1|\alpha, \alpha) = \sum_{i=1}^{c_\alpha} P_{\alpha i} p_{G2}(x_0, x_1; \mu_{\alpha i}, \sigma_{\alpha i}, \rho_\alpha) \quad (5.5)$$

όπου  $p_{G2}(x_0, x_1; \mu_{\alpha i}, \sigma_{\alpha i}, \rho_\alpha)$  είναι οι διοδιάστατες συναρτήσεις πυκνότητας πιθανότητας μορφής Gauss, με παραμέτρους  $(\mu_{\alpha i}, \mu_{\alpha i}, \sigma_{\alpha i}, \sigma_{\alpha i}, \rho_\alpha)$ .

Εναλλακτική και πιο αξιόπιστη προσέγγιση είναι η απευθείας χρήση διοδιάστατων κατανομών που περιγράφουν τις “προβληματικές” καταστάσεις. Έτσι παίρνοντας όλους τους δυνατούς συνδυασμούς μεταξύ των παρατηρούμενων πληθυσμών του στατικού και του κινούμενου τμήματος, καθώς και τις παραμέτρους τους και με την χρήση του προτεινόμενου εκτιμητή

Μέγιστης Πιθανοφάνειας (Παράρτημα Β) προκύπτουν επαναληπτικά οι βέλτιστες τιμές για τις παραμέτρους των δισδιάστατων κατανομών. Κατά την διάρκεια της επεξεργασίας αρκετές κλάσεις απορρίπτονται εξαιτίας της πολύ μικρής πιθανότητας που παρουσιάζουν. Το κυριότερο μειονέκτημα αυτής της προσέγγισης είναι το μεγάλο υπολογιστικό κόστος που απαιτείται, το οποίο είναι δυσανάλογο σε σχέση με την βελτίωση που αποφέρει στα εξαγόμενα αποτελέσματα.

### 5.3 Μοντελοποίηση Προβλήματος-Αποτελέσματα

Με βάση την προηγούμενη στατιστική ανάλυση είναι πλέον εφικτή η δημιουργία ενός όρου που εκφράζει την συνέπεια-συμβατότητα μεταξύ παρατηρήσεων και αποφάσεων. Χρησιμοποιώντας το ίδιο μοντέλο γειτονιάς με αυτό στην περίπτωση της ανίχνευσης μεταβολών, μπορούμε να μοντελοποιήσουμε στατιστικά το πρόβλημα με χρήση πεδίων Markov, και γειτονιά δεύτερης τάξης, όπου κατανομή Gibbs χρησιμοποιείται για την στατιστική περιγραφή της εξαγόμενης τμηματοποίησης. Ετσι η εκ των υστέρων πιθανότητα για την συνολική τμηματοποίηση  $\omega$  δίνεται από την σχέση:

$$p(\omega) = \frac{1}{Z} e^{-\frac{1}{T} U(I_t, I_{t+1}, \omega)} \quad (5.6)$$

με

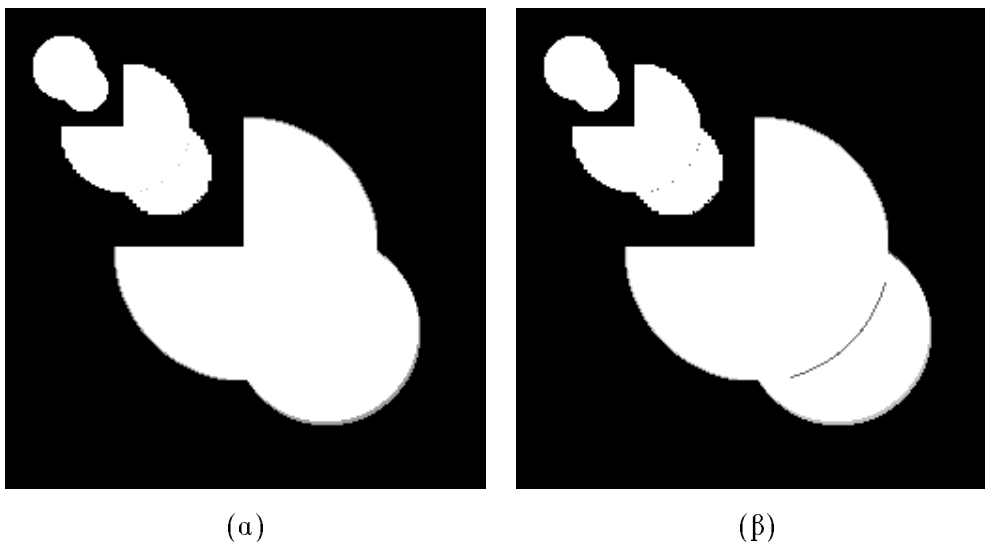
$$U(I_t, I_{t+1}, \omega) = U_1(\omega) + U_2(I_t, I_{t+1}, \omega) \quad (5.7)$$

όπου για τον υπολογισμό του όρου  $U_2(I_t, I_{t+1}, \omega)$  χρησιμοποιείται το αποτέλεσμα της ανίχνευσης κινούμενων αντικειμένων.

Η μορφή των όρων  $U_1$  και  $U_2$  είναι όμοια με την μορφή στο πρόβλημα της ανίχνευσης μεταβολών, ενώ ο ορισμός της συνάρτησης δυναμικού είναι πιο πολύπλοκος

$$V_c(\omega(i, j)) = \xi e_k^\top \begin{bmatrix} -\alpha_s & 1 & 1 & 1 \\ 1 & -\alpha_d & \alpha_{dd} & 1 \\ 1 & \alpha_{dd} & -\alpha_d & 1 \\ 1 & 1 & 1 & -\alpha_s \end{bmatrix} \begin{bmatrix} n_{BB} \\ n_{BO} \\ n_{OB} \\ n_{OO} \end{bmatrix} \quad (5.8)$$

όπου ισχύει η ακόλουθη σχέση  $\{(B, B) : 1, (B, O) : 2, (O, B) : 3, (O, O) : 4\}$  και  $n_{BB}$  (αντίστοιχα  $n_{BO}, n_{OB}, n_{OO}$ ) είναι ο αριθμός των σημείων στην γειτονιά του  $(i, j)$  με απόφαση  $(B, B)$ , και  $e_k$  μοναδιαίο διάνυσμα του  $R^4$  με την  $k$  συνιστώσα 1 και όλες τις υπόλοιπες 0. Η τιμή  $\alpha_s$  ενισχύει τις αποφάσεις  $(B, B)$  και  $(O, O)$ , το  $\alpha_d$  τις  $(B, O)$  και  $(O, B)$  όταν γειτονικά σημεία έχουν ίδιες αποφάσεις. Η τιμή  $\alpha_{dd}$  εκφράζει το κόστος να έχουμε γειτονικά σημεία με απόφαση  $(O, B)$  σε σημείο που έχει ήδη απόφαση  $(B, O)$  (ή το αντίθετο) ενώ σε άλλη περίπτωση το κόστος για διαφορετικές αποφάσεις σε γειτονικά σημεία είναι 1.0. Ο λόγος για τον οποίο χρησιμοποιείται η τιμή  $\alpha_{dd}$  είναι ότι τα γεγονότα  $(B, O)$  και  $(O, B)$  είναι αμοιβαία αποκλειόμενα για γειτονικά σημεία. Τέλος το  $\xi$  είναι το βάρος για την συνάρτηση δυναμικού, σε σχέση με τον όρο που εκφράζει συμβατότητα αποφάσεων-παρατηρήσεων.



Σχήμα 5.3: Ανίχνευση περιοχών κίνησης στην ακολουθία *Sphera*

(α) Πιθανοτική προσέγγιση με τμηματικά ομοιόμορφες κατανομές

(β) Πιθανοτική προσέγγιση με χρήση κατανομών Gauss

Η τελική τμηματοποίηση προκύπτει με την χρήση κριτηρίου μεγιστοποίησης της εκ των υστέρων πιθανότητας, το οποίο ισοδυναμεί με την ελαχιστοποίηση της ακόλουθης συνάρτησης κόστους:

$$U(I_t, I_{t+1}, \omega) = \sum_{c \in C_s} V_c(\omega) - \sum_{(i,j)} \ln[p(I_t, I_{t+1} | \omega(i, j))] \quad (5.9)$$

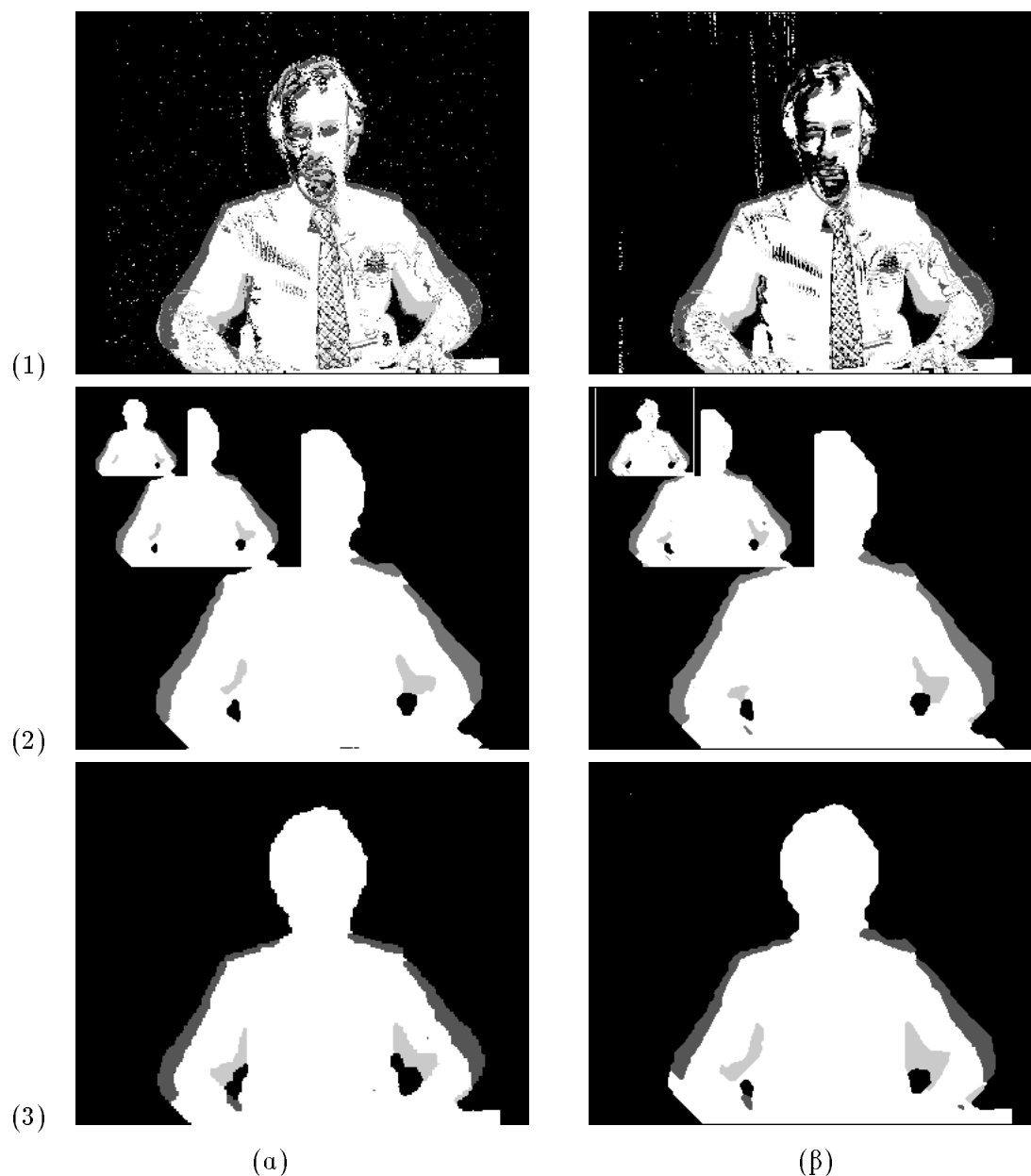
Για την ελαχιστοποίηση της παραπάνω συνάρτησης κόστους χρησιμοποιούνται οι μέθοδοι που εφαρμόστηκαν και στην ανίχνευση μεταβολών. Έτσι για την μείωση του υπολογιστικού κόστους, χρησιμοποιούνται τεχνικές πολλαπλών επιπέδων ενώ η σύγκλιση στην λύση σε κάθε επίπεδο γίνεται με αιτιοκρατικούς αλγόριθμους χαλάρωσης. Σημαντική παρατήρηση όσον αφορά την εκκίνηση των αλγορίθμων είναι ότι τα σημεία που ανήκουν στο στατικό τμήμα με βάση την λύση του προβλήματος ανίχνευσης μεταβολών αρχικοποιούνται με την απόφαση (B,B). Η αρχικοποίηση αυτή επιφέρει σημαντική μείωση του υπολογιστικού κόστους ιδιαίτερα για τον αλγόριθμο ΜΕκΠ.

Η μέθοδος δοκιμάστηκε αρχικά σε συνθετικά δεδομένα και πιο συγκεκριμένα στην ακολουθία *Sphera* όπου η λύση που προέκυψε ήταν αρκετά ικανοποιητική (Σχήμα 5.3). Εδώ πρέπει να σημειώσουμε ότι η ποιότητα της λύσης εξαρτάται σε πολύ μεγάλο βαθμό από την λύση του προβλήματος ανίχνευσης μεταβολών. Μια καλή λύση στο πρόβλημα αυτό οδηγεί στην δημιουργία ενός αξιόπιστου στατιστικού μοντέλου όσον αφορά το ζεύγος αποφάσεων ικανό στην συνέχεια με την χρήση του προτεινόμενου αλγορίθμου να μας δώσει τμηματοποίηση που δεν απέχει πολύ από την πραγματική λύση.

Στο Σχήμα (5.4) δίνονται τα αποτελέσματα της τμηματοποίησης για την ακολουθία

*Trevor White.* Για τον έλεγχο της αξιοπιστίας των μεθόδων καθορισμού της συνάρτησης πυκνότητας πιθανότητας των φτεινοτήτων ως προς κάποια απόφαση, παρατίθενται τα βέλτιστα αποτελέσματα σε κάθε περίπτωση. Επίσης υπάρχουν και οι τμηματοποιήσεις που θα προέκυπταν με την χρήση κριτηρίου μέγιστης πιθανοφάνειας. Το μαύρο αντιστοιχεί σε υπόστρωμα τόσο στην πρώτη όσο και στην δεύτερη εικόνα, οι δύο αποχρώσεις του γκρι σε επικαλυπτόμενα και αποκαλυπτόμενα τμήματα, ενώ τέλος το άσπρο σε κινούμενα αντικείμενο και στις δύο εικόνες. Επίσης η προβολή αυτών των αποτελεσμάτων στις δύο εικόνες εισόδου, και ο εντοπισμός της θέσης του κινούμενου αντικειμένου τις αντίστοιχες χρονικές στιγμές (Σχήμα 5.5) μας δίνει μία ιδέα για την αξιοπιστία και την ποιότητα της λύσης.





Σχήμα 5.4: Εντοπισμός Θέσης Κινούμενων Αντικειμένων για την ακολουθία *Trevor White*

(α) Ανάλυση ιστογράμματος με κατανομές Gauss

(β) Ανάλυση ιστογράμματος με τμηματικά ομοιόμορφες κατανομές

(1) Τμηματοποίηση με κριτήριο Μέγιστης Πιθανοφάνειας

(2) Τμηματοποίηση με χρήση Πολλαπλών επιπέδων και ΕΤΕΤ

(3) Τμηματοποίηση με χρήση Πολλαπλών επιπέδων και ΜΕκΠ



(α)



(β)

Σχήμα 5.5: Εντοπισμός θέσης Κινούμενων Αντικειμένων στο ζεύγος των εικόνων

## Κεφάλαιο 6

# Συμπεράσματα-Μελλοντική Εργασία

Στη εργασία αυτή προτείνονται τεχνικές και σχετιζόμενοι αλγόριθμοι για την επίλυση δύο βασικών προβλημάτων που εμφανίζονται στην ανίχνευση κίνησης. Τα προβλήματα αυτά είναι η ανίχνευση μεταβολών και ο εντοπισμός της θέσης των κινούμενων αντικειμένων.

Όσον αφορά το πρόβλημα της ανίχνευσης μεταβολών, η συνεισφορά αυτής της εργασίας στο θέμα είναι η ανάλυση του ιστογράμματος της διαφοράς φωτεινότητας μεταξύ των δύο εικόνων σαν μίξη κατανομών Laplace, με πολύ αποδοτικό τρόπο. Έτσι το κατώφλι για την απόφαση Μέγιστης Πιθανοφάνειας είναι δυναμικό ως προς τα δεδομένα. Επιπλέον η χρήση της προτεινόμενης μεθόδου για εκτίμηση της παραμέτρου εξομάλυνσης, έχει σαν αποτέλεσμα την δημιουργία ενός αυτόνομου μοντέλου όπου δεν απαιτείται καθόλου γνώση και συνεισφορά του χρήστη στη λύση του προβλήματος. Στην συνέχεια επιλέγεται μοντέλο Gibbs για την περιγραφή των αποφάσεων και τεχνικές πολλαπλών επιπέδων συνδυαζόμενες με αυτοκρατικούς αλγόριθμους χαλάρωσης για την επίλυση του προβλήματος. Ο συνδυασμός αυτός μας επιτρέπει την επίλυση του προβλήματος με μικρό υπολογιστικό κόστος. Επιπλέον το καλό στατιστικό μοντέλο που χρησιμοποιείται για να περιγράψει το πρόβλημα, έχει σαν αποτέλεσμα την εξαγωγή ικανοποιητικών αποτελεσμάτων ακόμα και στην περίπτωση κινούμενου παρατηρητή.

Η τμηματοποίηση της εικόνας σε στατικά και κινούμενα τμήματα χρησιμοποιείται στην συνέχεια για την εξαγωγή πληροφορίας που σχετίζεται με επικαλυπτόμενες και αποκαλυπτόμενες περιοχές του κινούμενου τμήματος. Σαν αποτέλεσμα καθορίζονται οι θέσεις των αντικειμένων που κινούνται και στις δύο εικόνες. Στο πρώτο βήμα επεξεργασίας η πυκνότητα πιθανότητας τόσο του στατικού, όσο και του κινούμενου τμήματος υπολογίζονται καθορίζοντας, είτε ένα δυναμικό μοντέλο μίξης κατανομών, είτε χρησιμοποιώντας τμηματικά ομοιόμορφες κατανομές με άμεσο στόχο την μείωση του απαιτούμενου κόστους. Τρεις

διαφορετικές προσεγγίσεις υιοθετούνται για την μοντελοποίηση και τον υπολογισμό της πυκνότητας πιθανότητας του ζεύγους φωτεινοτήτων στο ίδιο σημείο στα διαδοχικά στιγμιότυπα. Τα δύο πρώτα αποτελούν άμεση επέκταση της ανάλυσης μίξης των κατανομών στον μονοδιάστατο χώρο, ενώ το άλλο προκύπτει από την χρήση ομοιόμορφων κατανομών, όπως αυτές προκύπτουν με την τεχνική του κβαντισμού. Για το πρόβλημα της ανακατασκευής και σε αυτή την περίπτωση επιλέγεται μοντέλο Gibbs για την περιγραφή των αποφάσεων. Η λύση τελικά προκύπτει με χρήση τεχνικών πολλαπλών επιπέδων σε συνδυασμό με αιτιοκρατικούς αλγόριθμους χαλάρωσης. Ικανοποιητικά πειραματικά αποτελέσματα προέκυψαν σε συνθετικές και πραγματικές ακολουθίες εικόνων για εφαρμογές τηλεσυνεδριάσεων.

## 6.1 Συμπεράσματα

Για τον έλεγχο της αξιοπιστίας των προτεινόμενων μοντέλων, πειραματικά αποτελέσματα προέκυψαν από ακολουθίες εικόνων (πραγματικές ή συνθετικές) με διαφορετικό περιεχόμενο.

Οι μετρήσεις που αφορούν το υπολογιστικό κόστος έγιναν σε μηχάνημα Sun Station Sparc-10, ενώ η προτεινόμενη τεχνική εκτίμησης της παραμέτρου εξομάλυνσης δοκιμάστηκε για το πρόβλημα της ανίχνευσης μεταβολών. Τέλος τα συγκριτικά αποτελέσματα που αφορούν τους αιτιοκρατικούς αλγόριθμους χαλάρωσης προέρχονται από την ανίχνευση μεταβολών, ενώ τα συμπεράσματα για τις τεχνικές πολλαπλών επιπέδων προέρχονται από την λύση του δεύτερου προβλήματος, του εντοπισμού θέσης των κινούμενων αντικειμένων.

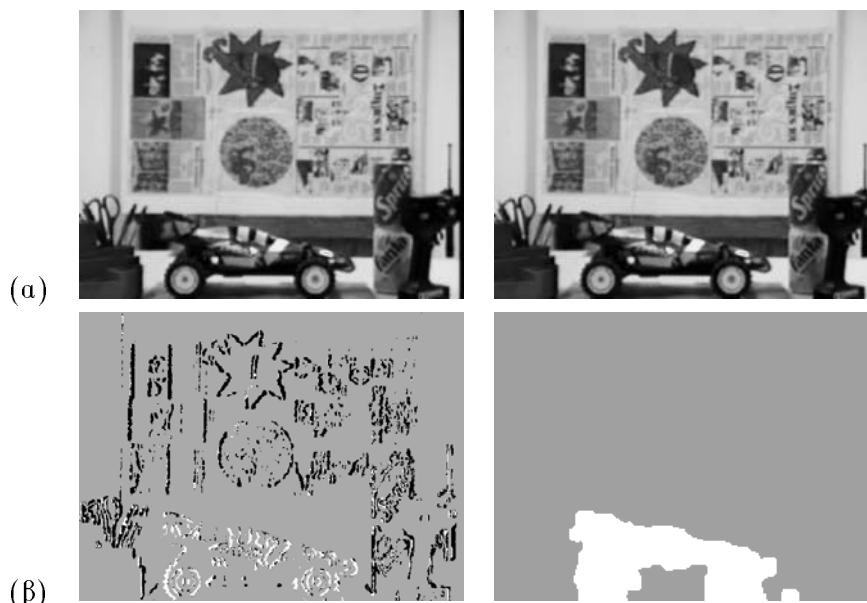
### 6.1.1 Σύγκριση με Άλλες Προσεγγίσεις

Για τον καλύτερο έλεγχο της προτεινόμενης μεθόδου, το τμήμα της εργασίας που αφορά την ανίχνευση ανεξάρτητα κινούμενων αντικειμένων, συγκρίνεται με το αντίστοιχο τμήμα της διδακτορικής διατριβής του Α. Αργυρού. Στην υλοποίηση αυτή έχουμε απευθείας υπολογισμό της 3-Δ κίνησης καθώς και του βάθους με την χρήση ζεύγους στερεοσκοπικών εικόνων. Η τελική τμηματοποίηση προκύπτει σαν συνδυασμός τόσο της εκτίμησης βάθους, όσο και της τρισδιάστατης κίνησης ([6]).

Στην περίπτωση αυτή έχουμε κίνηση του παρατηρητή, αλλά και ανεξάρτητα κινούμενο αντικείμενο. Ετσι αρχικά έχουμε εκτίμηση της κυρίαρχης κίνησης, και στην συνέχεια λύνεται το πρόβλημα της ανίχνευσης ανεξάρτητα κινούμενων αντικειμένων (Σχήμα 6.1, άσπρο χρώμα).

### 6.1.2 Τεχνική Ψηφοδοσίας

Η προτεινόμενη τεχνική για την βέλτιστη επιλογή της παραμέτρου εξομάλυνσης φαίνεται να δίνει ικανοποιητική λύση στο πρόβλημα. Η χρήση της σε συνδυασμό με την χρήση πολλαπλών

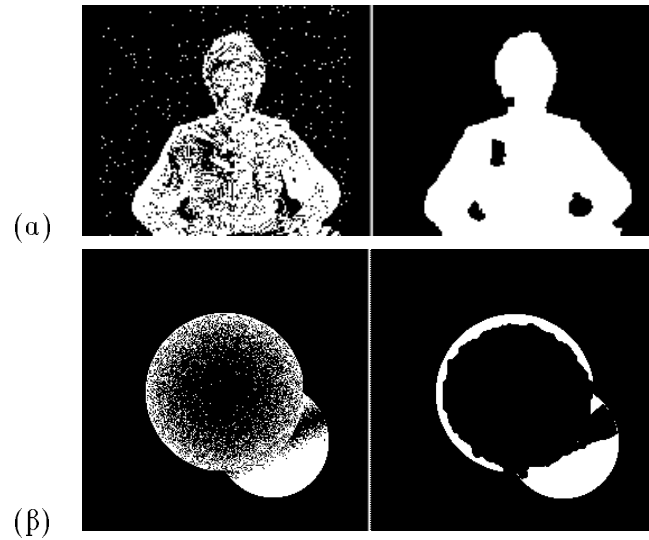


Σχήμα 6.1: Συγκριτικά Αποτελέσματα, Ακολουθία

(α) Εικόνες Ακολουθίας, (β) Ανίχνευση Ανεξάρτητα Κινούμενων Αντικειμένων: Με χρήση πεδίων Βάθους και 3-Δ κίνησης (αριστερά), Προτεινόμενη προσέγγιση (δεξιά)

επιπέδων όπου αρχικά το πλήθος των σημείων που ψηφίζουν είναι αρκετά μικρό, μπορεί να μας εξασφαλίσει μία αρκετά γρήγορη και καλή λύση. Όσον αφορά τα αποτελέσματα όπως αυτά εμφανίζονται στο τμήμα της ανίχνευσης μεταβολών είναι αρκετά ικανοποιητικά. Η αξιοπιστία της μεθόδου δοκιμάστηκε σε 6 διαφορετικές ακολουθίες εικόνων και σε πολλά διαδοχικά ζεύγη, και μπορούμε να πούμε ότι τα αποτελέσματα ήταν βέλτιστα στο μεγαλύτερο ποσοτό τους. Σαν συμπέρασμα από την χρήση της μεθόδου μπορούμε να πούμε ότι από τα πρώτα βήματα προκύπτει μία λύση πολύ κοντά στην βέλτιστη, γεγονός που μας δίνει την δυνατότητα εφαρμογής της μεθόδου μερικώς με άμεσο αποτέλεσμα την μείωση του απαιτούμενου υπολογιστικού κόστους.

Μη βέλτιστα αποτελέσματα προέκυψαν σε μερικά τμήματα δύο ακολουθιών. Είτοι ένα αρκετά ικανοποιητικό αποτέλεσμα αν και όχι βέλτιστο προκύπτει από την χρήση αυτής της μεθόδου στην ακολουθία *Trevor White* (Σχήμα 6.2(α)). Λαμβάνοντας υπόψη την λύση που προκύπτει με την χρήση κριτηρίου Μέγιστης Πιθανοφάνειας, μπορούμε να πούμε ότι το αποτέλεσμα είναι το καλύτερο δυνατό, αφού η συνάρτηση δυναμικού έχει απλοποιηθεί σε μεγάλο βαθμό (με την χρήση μόνο μίας παραμέτρου), οπότε μειώνεται σημαντικά και η ευελιξία του μοντέλου. Η μέθοδος φαίνεται να αποτυγχάνει για την ακολουθία *Sphera* αφού η λύση που προκύπτει είναι αρκετά μακριά από την επιθυμητή (Σχήμα 6.2(β)). Η φαινομενική “αποτυχία” της μεθόδου μπορεί να εξηγηθεί, αφού η αρχική ανακατασκευή (αριστερή εικόνα) δίνει πολύ συνεκτικό τμήμα για την “στατική” υπόθεση, γεγονός που λογικά οδηγεί σε μια μεγάλη τιμή της παραμέτρου εξομάλυνσης προς επιβεβαίωση αυτής της συνεκτικότητας, με αποτέλεσμα την



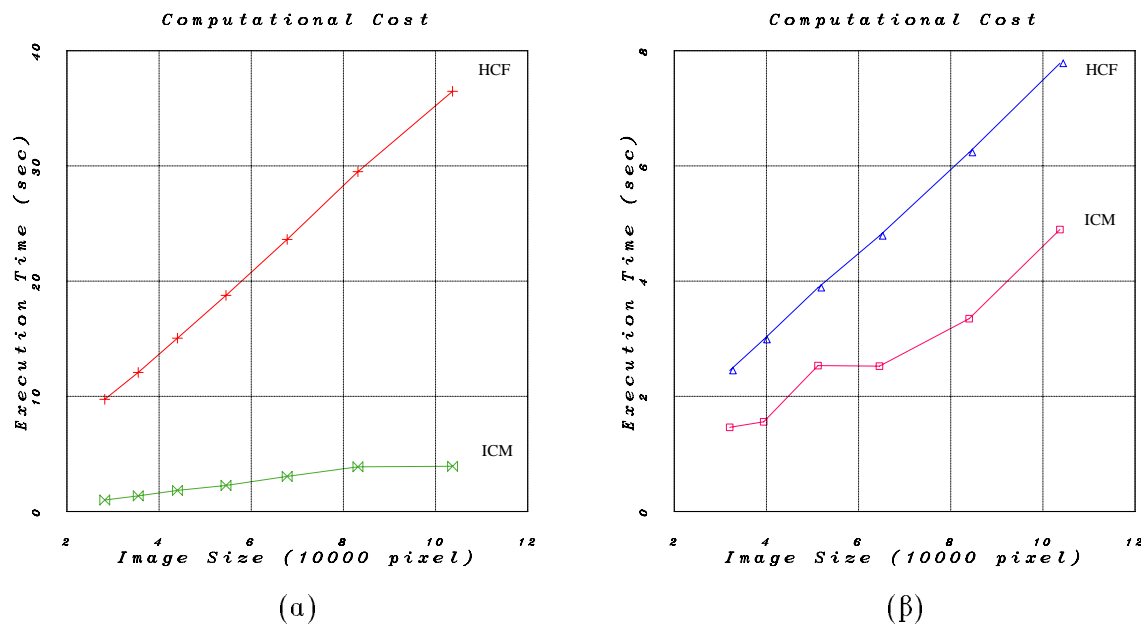
Σχήμα 6.2: Αυτόματος καθορισμός της παραμέτρου εξομάλυνσης με μη βέλτιστη ανακατασκευή: (α) Ακολουθία *Trevor White*, (β) Ακολουθία *Sphera*

αφαίρεση του μικρού απομονωμένου ποσοστού “κινούμενων” σημείων.

### 6.1.3 Αλγόριθμοι Ελαχιστοποίησης

Αν και οι δύο αλγόριθμοι που χρησιμοποιήθηκαν για την ελαχιστοποίηση των παραγόμενων συναρτήσεων κόστους ανήκουν στην ίδια κατηγορία, σημαντικές διαφορές παρατηρούνται όσον αφορά την συμπεριφορά τους. Τρία βασικά χαρακτηριστικά αποτελούν αντικείμενο μελέτης. Το υπολογιστικό κόστος, η ευαισθησία στην παράμετρο εξομάλυνσης και η εξάρτηση που προκύπτει με την χρήση κριτηρίου Μέγιστης Πιθανοφάνειας. Το τελευταίο θέμα είναι ισοδύναμο με την ανοχή του μοντέλου στον θόρυβο, ο οποίος κυρίως επιδρά στην αρχική λύση. Τα συμπεράσματα ακολουθούν.

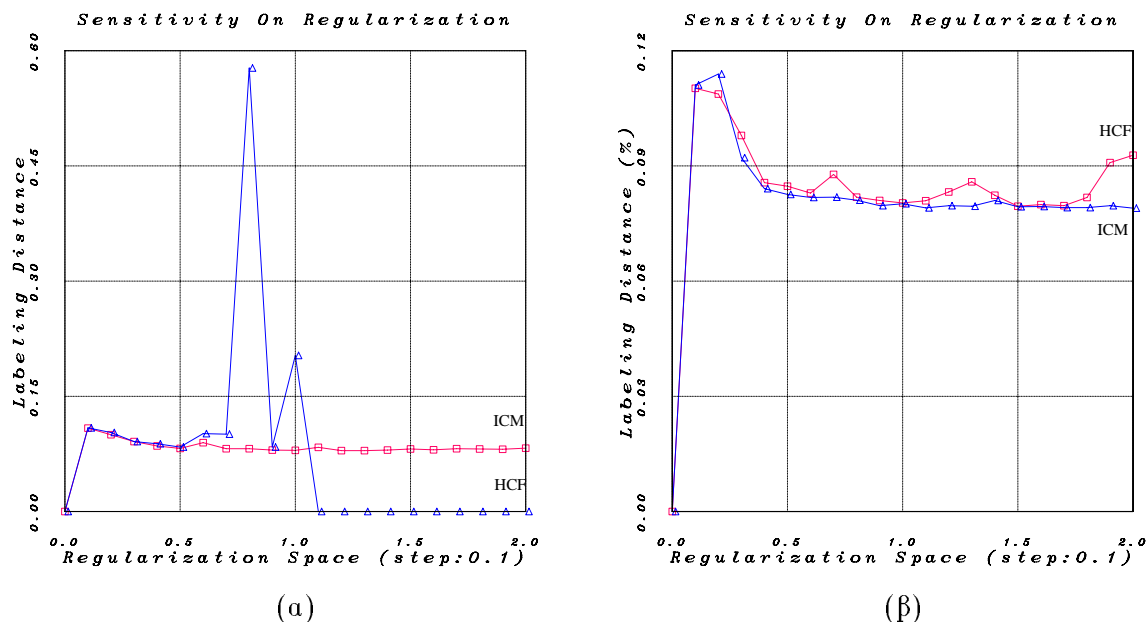
- **Υπολογιστικό κόστος:** Από τα αποτελέσματα και τις μετρήσεις που έγιναν είναι φανερό ότι ο αλγόριθμος ETET παρουσιάζει αρκετά μικρότερο κόστος από τον ΜΕκΠ. Το φαινόμενο αυτό είναι πολύ έντονο στην περίπτωση επεξεργασίας σε ένα επίπεδο, ενώ μειριάζεται κάπως με την χρήση τεχνικών πολλαπλών επιπέδων. Σημαντική επίσης παρατήρηση είναι η γραμμικότητα που φαίνεται να έχει το κόστος του ΜΕκΠ σε συνάρτηση με το μέγεθος της εικόνας, κάτι που θα περίμενε κανείς να ισχύει για τον ETET. Τέλος αξιοσημείωτο είναι το γεγονός ότι τα πολλαπλά επίπεδα δεν επιφέρουν σημαντική βελτίωση στο κόστος των ETET και πολλές φορές το κόστος είναι μεγαλύτερο από αυτό που προκύπτει χωρίς την χρήση τους. Αντιθέτως για τον ΜΕκΠ έχουμε σημαντική μείωση (τουλάχιστον 5 φορές), κάτι που μπορεί να ερμηνευθεί αφού το κόστος της δημιουργίας καθώς και της διαχείρισης του απαιτούμενο σωρού (τουλάχιστον στην



Σχήμα 6.3: Υπολογιστικό κόστος: (α) Ένα Επίπεδο, (β) Πολλαπλά Επίπεδα

αρχή) μειώνεται σημαντικά (Σχήμα 6.3).

- Ευαισθησία στην Παράμετρο Εξομάλυνσης:** Αρκετά διαφορετική είναι η συμπεριφορά των αλγορίθμων και σε αυτό το θέμα. Ο αλγόριθμος ΜΕΚΠ παρουσιάζει μεγάλη ευαισθησία στις τιμές του  $\lambda$  ιδιαίτερα στην περίπτωση επεξεργασίας μόνο σε ένα επίπεδο. Το πρόβλημα γίνεται εντονότερο στο διάστημα γύρω από την βέλτιστη τιμή, όπου πολύ μικρές μεταβολές της παραμέτρου εξομάλυνσης αποφέρουν σημαντικές αλλαγές στην ανακατασκευή. Το πρόβλημα κάπως μετριάζεται με την χρήση πολλαπλών επιπέδων. Αντίθετα ο αλγόριθμος ΕΤΕΤ φαίνεται να είναι αρκετά σταθερός αφού ακόμα και αρκετά μεγάλες διαφοροποιήσεις στην τιμή της παραμέτρου εξομάλυνσης δεν φαίνονται ικανές να προκαλέσουν ανάλογες μεταβολές στο αποτέλεσμα της ανακατασκευής. Η διαφοροποίηση αυτή μπορεί να εκληφθεί τόσο σαν μειονέκτημα, όσο και σαν πλεονέκτημα. Ετσι είναι λογικό να απαιτούμε ένα σταθερό μοντέλο όσον αφορά την συμπεριφορά ως προς τις παραμέτρους που το χαρακτηρίζουν, αλλά είναι επίσης επιθυμητό να έχουμε ένα ευέλικτο μοντέλο. Ίσως ο συνδυασμός των δύο παραπάνω αλγορίθμων θα μπορούσε να δώσει τον παραπάνω επιθυμητό συνδυασμό. Το Σχήμα 6.4 τεκμηριώνει τα παραπάνω συμπεράσματα. Ετσι τα διαγράμματα μας δίνουν την διαφοροποίηση της λύσης σε μικρές μεταβολές της παραμέτρου εξομάλυνσης. Το βήμα μεταβολής για την παράμετρο είναι 0.1 και στα διαγράμματα φαίνεται το ποσοστό των διαφορετικών αποφάσεων που προέκυψαν σε σχέση με τις αποφάσεις που είχαμε σε σχέση με την παράμετρο πριν γίνει η μεταβολή.

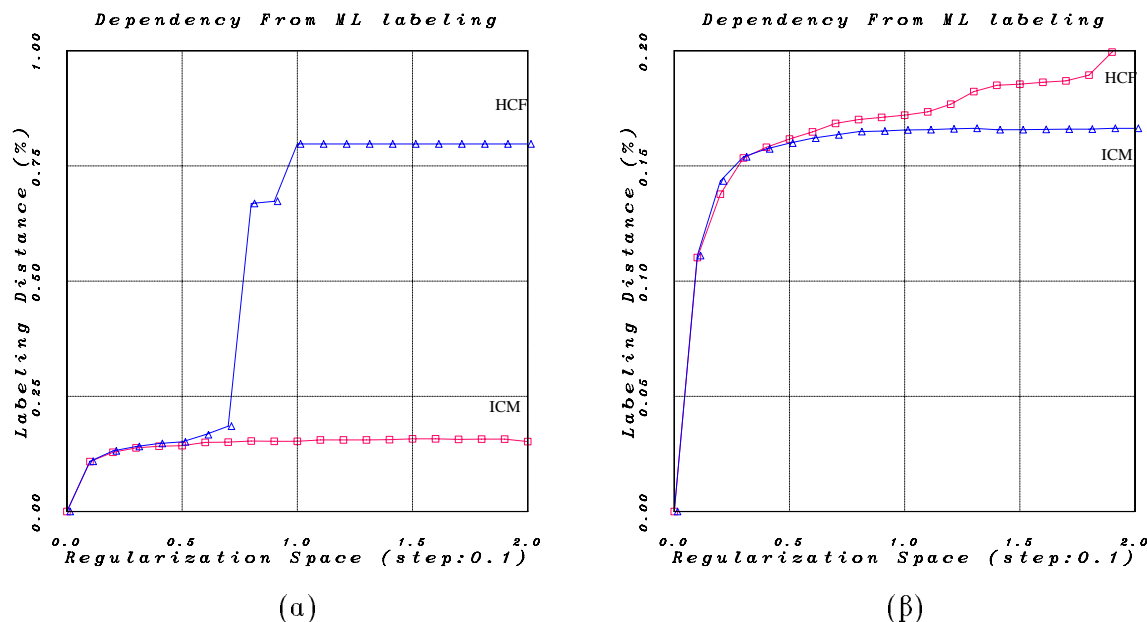


Σχήμα 6.4: Ευαισθησία στην Παράμετρο Εξομάλυνσης: (α) Ένα Επίπεδο, (β) Πολλαπλά Επίπεδα

- Εξάρτηση από την λύση Μέγιστης Πιθανοφάνειας:** Η παρουσία θορύβου στις διαθέσιμες παρατηρήσεις, πολλές φορές επιδρά σημαντικά στην ανακατασκευή ιδιαίτερα όταν αυτή προκύπτει με την χρήση κριτηρίου Μέγιστης Πιθανοφάνειας. Ο αλγόριθμος ΜΕκΠ με βάση τα πειραματικά δεδομένα, φαίνεται να μπορεί να αντιμετωπίζει αυτό το πρόβλημα, αφού ανάλογα με την τιμή της παραμέτρου εξομάλυνσης μπορεί να δώσει λύση αρκετά μακριά από την λύση Μέγιστης Πιθανοφάνειας. Βέβαια αυτή η δυνατότητα εμπεριέχει τον κίνδυνο της απόκλισης από την λύση, αφού αρχικά παίρνονται αποφάσεις με ημιτελή γειτονιά και μεγάλη τιμή της παραμέτρου εξομάλυνσης είναι δυνατόν να προκαλέσει την διάδοση των αρχικών αποφάσεων σε όλη την εικόνα. Αντίθετα ο αλγόριθμος ΕΤΕΤ παρουσιάζει σημαντική εξάρτηση από την αρχική λύση γεγονός που είναι αρνητικό στις περιπτώσεις ύπαρξης θορύβου. Αυτό είναι αναμενόμενο αφού όλες οι αποφάσεις παίρνονται με πλήρη γειτονιά όσον αφορά τις αποφάσεις (Σχήμα 6.5).

Ανακεφαλαιώνοντας αν και ο αλγόριθμος ΕΤΕΤ έχει αρκετά μικρότερο κόστος από τον ΜΕκΠ, δεν είναι αρκετά ευέλικτος και δεν μπορεί να αντιμετωπίσει περιπτώσεις με ισχυρή παρουσία θορύβου (όπως αυτή εμφανίζεται στην λύση Μέγιστης Πιθανοφάνειας που αποτελεί και την εκκίνηση του). Έτσι η χρήση του προτείνεται σε περιπτώσεις με χαμηλό ποσοστό θορύβου, αφού μπορεί να δώσει μία γρήγορη και καλή ανακατασκευή. Από την άλλη πλευρά ο αλγόριθμος ΜΕκΠ είναι πολύ πιο ευέλικτος κάτι που μπορεί να αντισταθμίσει το σημαντικό μειονέκτημα του υπολογιστικού κόστους που απαιτεί. Σε περιπτώσεις με μεγάλο ποσοστό θορύβου μπορεί





Σχήμα 6.5: Εξάρτηση από την Λύση Μέγιστης Πιθανοφάνειας: (α) Ένα Επίπεδο, (β) Πολλαπλά Επίπεδα

σαφέστατα να οδηγήσει σε πολύ καλύτερη ανακατασκευή απο αυτή που προκύπτει με την χρήση του ETEE.

#### 6.1.4 Τεχνικές Πολλαπλών Επιπέδων

Τα συμπεράσματα που εξάγονται από την χρήση των τεχνικών πολλαπλών επιπέδων είναι αρκετά ικανοποιητικά. Έτσι σε κάθε περίπτωση η μείωση του υπολογιστικού κόστους είναι τουλάχιστον 5 φορές, ενώ υπάρχουν περιπτώσεις που η μείωση αυτή είναι και 10 φορές σε σχέση με το κόστος σε ένα επίπεδο. Θα περίμενε κανείς η τεχνική πολλαπλών επιπέδων τόσο στο χώρο των αποφάσεων όσο και των παρατηρήσεων να έχει μεγαλύτερο κόστος (το κόστος δημιουργίας της Gauss πυραμίδας στο χώρο των παρατηρήσεων), αλλά κάτι τέτοιο δεν παρατηρείται έντονα. Με βάση τα παραπάνω μπορούμε να πούμε ότι το υπολογιστικό τους κόστος είναι ισοδύναμο.

Σημαντική διαφορά υπάρχει στην πολυπλοκότητα των μοντέλων. Έτσι για τα πολλαπλά επίπεδα στο χώρο δεδομένων και αποφάσεων, απαιτούνται διαφορετικές τιμές για τις παραμέτρους στην συνάρτηση κόστους στα διάφορα επίπεδα, καθώς και στην συνάρτηση των δυναμικών. Το πρόβλημα γίνεται εντονότερο κατά την χρήση του αλγόριθμου ΜΕΚΠ, που είναι αρκετά ευαίσθητος στις παραμέτρους. Αντίθετα για τα πολλαπλά επίπεδα μόνο στο χώρο των αποφάσεων, όπου η συνάρτηση κόστους έχει την ίδια μορφή αλλά και ίδιο πλήθος όρων σε όλα τα επίπεδα, μπορούν να χρησιμοποιηθούν οι ίδιες παράμετροι σε όλα τα επίπεδα. Ταυτόχρονα όμως η χρήση διαφορετικών παραμέτρων στα διάφορα επίπεδα μας δίνει την

δυνατότητα ευελιξίας όσον αφορά την ανακατασκευή. Έτσι μπορούμε με την χρήση κατάλληλων παραμέτρων στο αρχικό επίπεδο να πάρουμε ένα υπερούλο της αναμενόμενης λύσης, το οποίο στην συνέχεια με λείανση να δώσει την σωστή ανακατασκευή. Τέλος αν λάβουμε υπόψη ότι μία κακή λύση σε αρχικά επίπεδο είναι πολύ δύσκολο να διορθωθεί στην συνέχεια με την χρήση ίδιων παραμέτρων (αλλά ακόμα και διαφορετικών), μπορούμε να πούμε ότι τα πολλαπλά επίπεδα στο χώρο αποφάσεων και δεδομένων διαθέτουν μεγαλύτερη ευελιξία.

Τέλος διαφορετική συμπεριφορά υπάρχει και στην σταθερότητα στον θόρυβο. Η χρήση της ίδιας συνάρτησης κόστους σε όλα τα επίπεδα όπου η συμβατότητα μεταξύ αποφάσεων και παρατηρήσεων, γίνεται με την χρήση μπλόκ σημείων (όπου οι πιθανότητες προκύπτουν σαν άθροισμα πιθανοτήτων πάνω σε όλα τα σημεία του μπλοκ) είναι λογικό να δίνει πιο σταθερά αποτελέσματα και να επηρεάζεται λιγότερο από την παρουσία θόρυβο. Αντίθετα παρά την χρήση πυραμίδων Gauss για την δημιουργία της πυραμίδας στο χώρο των δεδομένων (η οποία κάνει και κατά κάποιο τρόπο λείανση σε αυτά) παρουσιάζονται λάθη στον υπολογισμό των συναρτήσεων πυκνότητας πιθανότητας μεταξύ δεδομένων και αποφάσεων. Αυτό γίνεται γιατί οι συναρτήσεις αυτές υπολογίζονται στο πλήρη χώρο ανάλυσης όπου έχουμε μεγαλύτερο και πιο αξιόπιστο στατιστικό δείγμα.

## 6.2 Μελλοντική Εργασία

Όσον αφορά μελλοντική εργασία καθώς και επεκτάσεις των μοντέλων, υπάρχουν ανοικτά θέματα που θα μπορούσαν να βελτιώσουν τα εξαγόμενα αποτελέσματα.

Σημαντική βελτίωση για τις περιπτώσεις κινούμενου παρατηρητή θα είχε ένας πιο σταθερός και πιο ακριβής αλγόριθμος υπολογισμού της κυρίαρχης κίνησης. Έτσι για τις περιπτώσεις που το αφηνικό μοντέλο δεν επαρκεί (κάτι που ισχύει για πολλά είδη κινήσεων), ο άμεσος υπολογισμός της κίνησης στον 3-Δ χώρο θα έδινε πολύ καλύτερα αποτελέσματα.

Ανοικτό είναι και το θέμα της χρήσης στοχαστικών αλγόριθμων χαλάρωσης για την ελαχιστοποίηση, αν και με βάση τη σχέση κόστους-απόδοσης που παρουσιάζουν οι αυτοκρατικοί αλγόριθμοι που έχουν υλοποιηθεί, είναι δύσκολο να προκύψουν αποτελέσματα με καλύτερη σχέση μεταξύ αυτών των σημαντικών παραμέτρων. Ιδιαίτερο ενδιαφέρον θα είχε και η παραλληλοποίηση του προβλήματος σε συνδυασμό με τις υπάρχουσες τεχνικές, με στόχο την επιπλέον μείωση του κόστους. Οι ιδιότητες των πεδίων Markov μας δίνουν άμεσα την δυνατότητα για κάτι τέτοιο, χωρίς ιδιαίτερες δυσκολίες.

Άμεση επέκταση της προτεινόμενη τεχνική καθορισμού της παραμέτρου εξομάλυνσης, θα ήταν η στατιστική οριοθέτηση της, αν αυτό είναι εφικτό. Κριτήριο στην περίπτωση αυτή θα ήταν η ελαχιστοποίηση μέσου τετραγωνικού σφάλματος, προερχόμενου από την απόσταση μεταξύ της ανακατασκευής και των διαθέσιμων παρατηρήσεων.

Τέλος σε σχέση με το πρόβλημα του εντοπισμού της θέσης των αντικειμένων, ανοιχτό είναι το πρόβλημα της χρήσης ενός εναλλακτικού σχήματος για την μοντελοποίηση του όρου που δίνει μέτρο ποιότητας της ανακατασκευής, σε σχέση με τις διαθέσιμες παρατηρήσεις. Ετοι για τις περιπτώσεις που το κινούμενο τμήμα στις εικόνες είναι σημαντικά μικρότερο από το στατικό, η στατιστική μοντελοποίηση θα μπορούσε να περιοριστεί τοπικά και να γίνεται για κάθε αντικείμενο ξεχωριστά σε ένα μικρό χώρο που το εμπεριέχει. Επιπλέον η χρήση άλλων διαθέσιμων πληροφοριών, όπως η κίνηση των περιγραμμάτων, η ανίχνευση ακμών και πληροφορίες σχετικά με την υφή των αντικειμένων, ίσως να βελτιώναν το αποτέλεσμα, με ορατό πάντα τον κίνδυνο για σημαντική αύξηση της πολυπλοκότητας του μοντέλου και μείωση της ευελιξίας του. Άμεση επέκταση της εργασίας αυτής και για ακολουθίες έγχρωμων εικόνων θα είχε επίσης ενδιαφέρον.



## Παράρτημα Α

# Κβαντιστής Lloyd-Max

Ενας κβαντιστής αντιστοιχεί σε μία συνεχή τυχαία μεταβλητή μια από τις διακριτές τιμές ενός πεπερασμένου συνόλου  $\{r_1, r_2, \dots, r_N\}$  αντιπροσωπευτικών τιμών. Για τον προσδιορισμό του απαιτούνται το σύνολο των αντιπροσωπευτικών τιμών και το σύνολο των σημείων απόφασης  $\{d_1, d_2, \dots, d_{N-1}\}$  που χωρίζουν τον άξονα των πραγματικών αριθμών στα ακόλουθα διαστήματα

$$R = (-\infty, d_1] \cup (d_1, d_2] \cup \dots \cup (d_{N-1}, \infty) \quad (\text{A.1})$$

Εστω  $X$  τυχαία μεταβλητή, της οποίας η συνάρτηση πυκνότητας πιθανότητας είναι  $p(x)$ . Ζητείται ο προσδιορισμός τόσο των επιπέδων κβαντισμού, όσο και των σημείων απόφασης για δοσμένο αριθμό  $N$  διαστημάτων και επιπέδων κβαντισμού, έτσι ώστε να ελαχιστοποιείται το μέσο τετραγωνικό σφάλμα

$$D = E\{(X - Q(X))^2\} = \sum_{i=1}^N \int_{d_{i-1}}^{d_i} (x - r_i)^2 p(x) dx \quad (\text{A.2})$$

όπου  $Q(\cdot)$  είναι ο τελεστής κβαντισμού και  $d_0 = -\infty, d_N = \infty$ .

Οι Lloyd-Max έδωσαν τις αναγκαίες συνθήκες για την ελαχιστοποίηση του  $D$ , που συνίστανται σε δύο συστήματα εξισώσεων. Το πρώτο δίνει την καλύτερη αντιπροσώπευση, με δοσμένα τα διαστήματα

$$r_i = E\{X|X \in (d_{i-1}, d_i)\} = \frac{\int_{d_{i-1}}^{d_i} x p(x) dx}{\int_{d_{i-1}}^{d_i} p(x) dx}, \quad i = 1, \dots, N \quad (\text{A.3})$$

Το δεύτερο σύστημα δίνει την καλύτερη τμηματοποίηση του άξονα των πραγματικών αριθμών με δοσμένα τα επίπεδα αντιπροσώπευσης

$$d_i = \frac{r_i + r_{i+1}}{2}, \quad i = 1, \dots, N - 1 \quad (\text{A.4})$$

Στην γενική περίπτωση το παραπάνω σύστημα των εξισώσεων (A.3) και (A.4) μπορεί να λυθεί μόνο με αριθμητικές επαναληπτικές μεθόδους. Μετά από κάποιες αρχικές τιμές για

τα επίπεδα κβαντισμού, οι παραπάνω εξισώσεις χρησιμοποιούνται διαδοχικά και επαναληπτικά μέχρι το μοντέλο να συγκλίνει. Πρέπει να σημειωθεί ότι η επαναληπτική επίλυση του προηγούμενου συστήματος των εξισώσεων δεν εξασφαλίζει πάντοτε την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος και μπορεί να δώσει λύση που αντιστοιχεί σε τοπικό ελάχιστο. Οι παραπάνω συνθήκες είναι σε κάθε περίπτωση αναγκαίες αλλά όχι ικανές να μας δώσουν την βέλτιστη λύση, ελαχιστοποιώντας την ποσότητα  $D$ .

Τέλος κριτήριο τερματισμού του αλγορίθμου που προκύπτει με την χρήση των παραπάνω εξισώσεων, αποτελεί η διόρθωση που επιτυγχάνεται από βήμα σε βήμα στο μέσο τετραγωνικό σφάλμα. Όταν η διαφορά μεταξύ των μέσων τετραγωνικών σφαλμάτων δύο διαδοχικών περασμάτων είναι κάτω από κάποιο εμπειρικά καθοριζόμενο κατώφλι, τότε ο αλγόριθμος θεωρείται ότι έχει συγκλίνει.

## Παράρτημα Β

# Εκτιμητής παραμέτρων σε προβλήματα Μίξης Κατανομών

Ένα πολύ συνηθισμένο πρόβλημα στην στατιστική, είναι η ανάλυση κάποιου στατιστικού δείγματος στις κλάσεις που το συνθέτουν, των οποίων το πλήθος θεωρείται γνωστό. Η υπόθεση που χρησιμοποιείται, είναι ότι το είδος της συνάρτησης πυκνότητας πιθανότητας που περιγράφει τις κλάσεις που συνθέτουν το παρατηρούμενο δείγμα, είναι γνωστό εκτός από τις παραμέτρους των για κάθε κλάση. Εστω  $c$  το πλήθος των κλάσεων (το οποίο γνωρίζουμε),  $P_j$  η εκ των προτέρων πιθανότητα της κλάσης  $j$ , και  $p(x|\theta_j)$  η συνάρτηση πυκνότητας πιθανότητας για την δοσμένη κλάση, όπου το διάνυσμα  $\theta_j$  αποτελεί το σύνολο των άγνωστων παραμέτρων για την  $p(x|\theta_j)$ . Η μίξη  $c$  κλάσεων δίνει την ακόλουθη συνάρτηση πυκνότητας πιθανότητας

$$p(x|\phi) = \sum_{j=1}^c P_j p(x|\theta_j), \quad \sum_{j=1}^c P_j = 1 \quad (\text{B.1})$$

όπου το διάνυσμα  $\phi$  αποτελείται από όλες τις άγνωστες παραμέτρους της μίξης, δηλαδή  $\{\theta_j : j = 1, \dots, c\}$  και  $\{P_j : j = 1, \dots, c\}$ . Το πρόβλημα είναι ισοδύναμο με τον υπολογισμό των άγνωστων παραμέτρων του διανύσματος  $\phi$ . Ο υπολογισμός αυτός μπορεί να γίνει με την χρήση κριτηρίου Μέγιστης Πιθανοφάνειας ([31], [45], και [52]). Ένας εναλλακτικός τρόπος για τον υπολογισμό των παραπάνω παραμέτρων είναι η χρήση επεξεργασίας fuzzy ISODATA ([8]). Στην εργασία αυτή χρησιμοποιήθηκε ο εκτιμητής Μέγιστης Πιθανοφάνειας, τον οποίου θα γίνει συνοπτική περιγραφή όσον αφορά τις συνθήκες που τον διέπουν καθώς και των εφαρμογών τους για κατανομές Gauss και Laplace που χρησιμοποιήθηκαν.

Η εκ των υστέρων πιθανότητα της κλάσης  $i$  δεδομένου της παρατήρησης  $x$  δίνεται από την σχέση

$$P_i(x|\phi) = \frac{P_i p(x|\theta_i)}{\sum_{j=1}^c P_j p(x|\theta_j)} \quad (\text{B.2})$$

Αν  $\{x_1, \dots, x_n, \dots, x_N\}$  είναι το σύνολο των δειγμάτων, τότε οι εκ των προτέρων πιθανότητες

καθώς και οι παράμετροι των συναρτήσεων πυκνότητας πιθανότητας του μοντέλου πρέπει να ικανοποιούν τις ακόλουθες εξισώσεις:

•

$$\hat{P}_i = \frac{1}{N} \sum_{n=1}^N \hat{P}_i(x_n|\phi) \quad (\text{B.3})$$

•

$$\sum_{n=1}^N \hat{P}_i(x_n|\phi) \nabla_{\theta_i} \log p(x_n|\hat{\theta}_i) = 0 \quad (\text{B.4})$$

όπου

$$\hat{P}_i(x_n|\hat{\phi}) = \frac{\hat{P}_i p(x_n|\hat{\theta}_i)}{\sum_{j=1}^c \hat{P}_j p(x_n|\hat{\theta}_j)} \quad (\text{B.5})$$

Στην περίπτωση μίξης κατανομών Laplace, ο ακόλουθος επαναληπτικός αλγόριθμος χρησιμοποιείται για τις παράμετρους  $\lambda_i$ :

$$\lambda_i(j+1) = \frac{N \hat{P}_i(j)}{\sum_{n=1}^N \hat{P}_i(x_n|\hat{\phi}(j)) |x_n|} \quad (\text{B.6})$$

ενώ για την περίπτωση μίξης κατανομών Gauss οι σχέσεις μας δίνουν τις μέσες τιμές και τις διασπορές σε κάθε βήμα είναι οι ακόλουθες

•

$$\mu_i(j+1) = \frac{\sum_{k=1}^n P(\omega_i|x_k, \hat{\phi}(j)) x_k}{\sum_{k=1}^n P(\omega_i|x_k, \hat{\phi}(j))} \quad (\text{B.7})$$

•

$$\sigma_i^2(j+1) = \frac{\sum_{n=1}^N P(\omega_i|x_k, \hat{\phi}(j)) (x_k - \hat{\mu}_i)^2}{\sum_{n=1}^N P(\omega_i|x_k, \hat{\phi}(j))} \quad (\text{B.8})$$

όπου  $\phi_i = (P_i, \mu_i, \sigma_i)$ .



# Βιβλιογραφία

- [1] T. Aach and A. Kaup. Bayesian algorithms for adaptive change detection in image sequences using markov random fields. *Signal Processing: Image Communication*, 7:147–160, 1995.
- [2] T. Aach, A. Kaup, and R. Mester. Statistical model-based change detection in moving video. *Signal Processing*, 31:165–180, 1993.
- [3] P. Anandan. Computing dense displacement fields with confidence measurement of visual motion. *Proceedings in SPIE Conference on Intelligent Robots and Computer Vision*, pages 184–194, 1984.
- [4] H. Andrews and B. Hunt. *Digital Image Restoration*. New Jersey: Prentice Hall, 1977.
- [5] G. Archer and D. Titterington. On some bayesian/regularization methods for image restoration. *IEEE Transactions on Image Processing*, 4:989–995, July 1995.
- [6] A. Argyros, M. Lourakis, P. Trahanias, and S. Orphanoudakis. Independent 3d motion detection through robust regression in depth layers. *FORTH Technical Report 159*, 1996.
- [7] S. Barnard. Stochastic stereo matching over scale. *International Journal of Computer Vision*, 3:17–32, July 1989.
- [8] J. Bedzek and J. Dunn. Optimal fuzzy partitions : A heuristic for estimating the parameters in a mixture of normal distributions. *IEEE Transactions on Computers*, pages 835–839, August 1972.
- [9] M. Bertero, T. Poggio, and V. Torre. Ill-posed problems in early vision. *Proceedings of the IEEE*, 76:869–889, August 1988.
- [10] J. Besag. Spatial interaction and the statistical analysis of lattice systems (with discussion). *Journal of Royal Statistics Society*, B:36:192–326, 1974.

- [11] J. Besag. On the statistical analysis of dirty images. *Journal of Royal Statistics Society*, 48:259–302, 1986.
- [12] Martin Bischel. Segmenting simply connected moving objects in a static scene. *IEEE Transactions on Pattern Analysis Machine Intelligence*, pages 1138–1142, November 1994.
- [13] C. Bouman and B. Liu. Multiple resolution segmentation of textured images. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 13:99–113, 1991.
- [14] P. Bouthemy and E. Francois. Motion segmentation and qualitative dynamic scene analysis from an image sequence. *International Journal of Computer Vision*, 10:157–182, 1993.
- [15] P. Bouthemy and P. Lalande. Motion detection in an image sequence using gibbs distribution. *Proceedings in ICASSP*, pages 1651–1664, May 1989 (Glasgow).
- [16] P. Bouthemy and P. Lalande. Detection and tracking of moving objects based on a statistical regularization method in space and time. *Proceedings in ECCV*, pages 307–311, 1990 (Antibes).
- [17] P. Bouthemy and P. Lalande. Recovery of moving object masks in an image sequence using spatiotemporal contextual information. *Optical Engineering*, June 1993.
- [18] P. Bouthemy and J-M. Odobez. Detection of multiple moving objects using multiscale mrf with camera motion compensation. *Proceedings in ICIP*, November 1994 (Austin).
- [19] C. Cafforio and F. Rocca. Methods for measuring small displacements of television images. *IEEE Transactions on Information Theory*, IT-22:1973–1979, 1976.
- [20] P. Chou and C. Brown. The theory and practice of bayesian image labeling. *International Journal of Computer Vision*, 4:185–210, 1990.
- [21] D. Connor, B. Haskell, and F. Mounts. A frame-to-frame picturephone coder for signals containing differential quantizing noise. *Bell System Technical Report*, 52:35–51, 1973.
- [22] G. Cross and A. Jain. Markov random fields texture models. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 5:25–39, 1983.
- [23] P. Davies. Aspects of robust linear regression. *The Annals of Statistics*, 21:1843–1889, 1993.
- [24] H. Derin and P. Kelly. Discrete-index markov-type random processes. *Proceedings of the IEEE*, 77:1485–1510, October 1989.

- [25] H. Derrin and H. Elliot. Modelling and segmentation of noisy and textured images using gibbs random fields. *IEEE Transactions on Pattern Analysis Machine Intelligence*, pages 39–55, 1987.
- [26] N. Diehl. Object-oriented motion estimation and segmentation in image sequences. *IEEE Transactions on Image Processing*, 3:1901–1904, February 1990.
- [27] J. Dinten, X. Guyon, and J. Yao. On the choise of the regularization parameter: the case of binary images in the bayesian restoration framework. *Spatial Statistics and Imaging (Editor: A. Posolo)*, pages 55–77, 1991.
- [28] G. Donohoe, D. Hush, and N. Ahment. Change detection for target detection and classification in video sequences. *Proceedings in ICASSP*, 88:1084–1087, April (New York) 1988.
- [29] J. Driessen, J. Biemond, and D. Boekee. A pel-recursive segmentation algoritmh for motion compensated image sequence coding. *Proceedings in ICASSP*, pages 1901–1904, April 1989 (New York).
- [30] R. Dubes and A. K. Jain. Random field models in image analysis. *Journal of Applied Statistics*, 16:131–164, 1989.
- [31] R. Duda and P. Hart. *Pattern Classification and Scene Analysis*. New York: Willey-Interscience, 1973.
- [32] W. Enkelman. Investigations of multigrid algorithms for the estimation of optical flow field in image sequences. *Computer Vision Graphics, Image Processing*, 43:150–177, 1988.
- [33] N. Galatsanos and A. Katsaggelos. Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. *IEEE Transactions on Image Processing*, 1:322–336, 1992.
- [34] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 6:721–741, 1984.
- [35] A. Gersho and R. Gray. *Vector quantization and signal processing*. Kluwer Publishers, 1990.
- [36] F. Heitz, P. Pérez, and P. Bouthemy. Multiscale minimization of global energy functions in some visual recovery problems. *CVGIP: Image Understanding*, 59:125–134, 1994.

- [37] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:195–203, 1981.
- [38] P. Huber. *Robust Statistics*. Wiley, 1981.
- [39] K. Karmann, A. Brandt, and R. Gerl. Moving object segmentation based on adaptive reference images. *Signal Processing: Theories and Applications*, V:951–954, 1990.
- [40] A. Kuria'nski and M. Nieniewski. Hidden mrf detection of motion of objects with uniform brightness. *Proc. on ICIAF*, pages 657–662, 1995 (San Remo).
- [41] P. Lalande and P. Bouthemy. A statistical approach to the detection and tracking of moving objects in an image sequence. *Signal Processing*, V:Theories and Applications:1947–1950, 1990.
- [42] J. Letag, V. Rebuffel, and P. Bouthemy. Motion detection robust to perturbations: a statistical regularization and temporal integration framework. *Proceedings in ICCV*, pages 21–30, 1993 (Berlin).
- [43] J. Limb, R. Pease, and K. Walsh. Combining interframe and frame-to-frame coder for television. *Bell System Technical Report*, 53:1137–1173, 1974.
- [44] A. Little and B. Rubin. On jointly estimating parameters and missing values by maximizing the complete data likelihood. *American Statistics*, 37:218–220, 1983.
- [45] G. McLachlan, D. Peel, and W. Whiten. Maximum likelihood clustering via normal mixture model. *Signal Processing: Image Communication*, 8:105–111, 1996.
- [46] L. Morri, F. Rocca, and S. Tubaro. Motion compensated interpolation using foreground/background segmentation. *Digital Signal Processing*, 37:379–384, 1991.
- [47] J-M. Odobez and P. Bouthemy. Robust multiresolution estimation of parametric motion models. *Visual Communication and Image Representation*, pages 348–365, December 1995.
- [48] N. Paragios, P. Perez, G. Tziritas, C. Labit, and P. Bouthemy. Adaptive detection of moving objects, using multiscale techniques. *To appear in ICIP*, 1996 (Laussane).
- [49] N. Paragios and G. Tziritas. Detection and location of moving objects, using deterministic relaxation algorithms. *To appear in ICPR*, 1996 (Vienne).
- [50] P. Pérez and F. Heitz. Restriction of a markov random field on a graph and multiresolution statistical image modeling. *IEEE Transactions on Information Theory*, 42:180–190, 1996.

- [51] A. Rosenfeld. *Multilevel relaxation in low-level computer vision*. Springer-Verlag, 1984.
- [52] R. Schalkoff. *Pattern Recognition : Statistical, Structural and Neural Approaches*. J. Wiley and sons, 1992.
- [53] B. Silverman. A fast efficient cross-validation method for smoothing parameter choice in spline regression. *Journal of the American Statistical Association*, pages 584–589, 1984.
- [54] Z. Sivan and D. Malah. Change detection and texture analysis for image sequence coding. *Signal Processing: Image Communication*, 6:357–376, 1994.
- [55] M. Stone. Cross-validatory choice and assessment of statistical prediction. *Journal of the Royal Statistical Society*, pages 129–139, 1974.
- [56] D. Terzopoulos. Image analysis using multigrid relaxation methods. *IEEE Transactions on Pattern Analysis Machine Intelligence*, pages 129–139, 1986.
- [57] R. Thoma and M. Bierling. Motion compensating interpolation considering covered and uncovered background. *Signal Processing: Image Communication*, 1:191–212, October 1989.
- [58] A. Tikhonov and V. Arsenin. *Solution of Ill-Posed Problems*. New York: J. Wiley, 1977.
- [59] D. Titterington. Choosing the regularization parameter in image restoration. *Spatial Statistics and Imaging (Editor: A. Posolo)*, pages 393–402, 1991.
- [60] G. Tziritas and C. Labit. *Motion Analysis for image sequence coding*. Elsevier, 1994.
- [61] O. Wenstop. Motion detection from image information. *Proceedings in Scandianvian Conference on Image Analysis*, pages 381–386, 1983 (Copenhagen).
- [62] S. Yu and M. Berthod. A game strategy for image labeling. *Computer Vision and Image Understanding*, 61:32–37, January 1995.



# Ευρετήριο Ορων

Απιοκρατικοί Αλγόριθμοι,	12
Αμφίδρομη Πυραμίδα Επεξεργασίας,	09
Ανίχνευση Μεταβολών,	25
Αφφινικό Μοντέλο,	28
Διαφορά Φωτεινότητων,	30
Διευρυμένη Τοπική Ενέργεια,	15
Επαναλαμβανόμενες Τοπικά Επικρατούσες Τιμές,	13
Ιδία Κίνηση,	27
Κατανομή Gibbs,	07
Μέγιστη Εμπιστοσύνη κατά Προτεραιότητα,	14
Μίξη Κατανομών,	31
Μονόδρομη Πυραμίδα Επεξεργασίας,	09
Παράμετρος Εξομάλυνσης,	19
Πεδία Markov,	05
Στοχαστικοί Αλγόριθμοι,	08
Συνάρτηση Δυναμικών,	06
Συνάρτηση Κόστους (Ενέργειας),	06
Συνάρτηση Προσαρμογής,	28
Σύστημα Γειννίαςας,	05
Τεχνικές Πολλαπλών Επιπέδων,	08
Τεχνική Εξομάλυνσης,	19
Χάρτης Αποφάσεων,	05