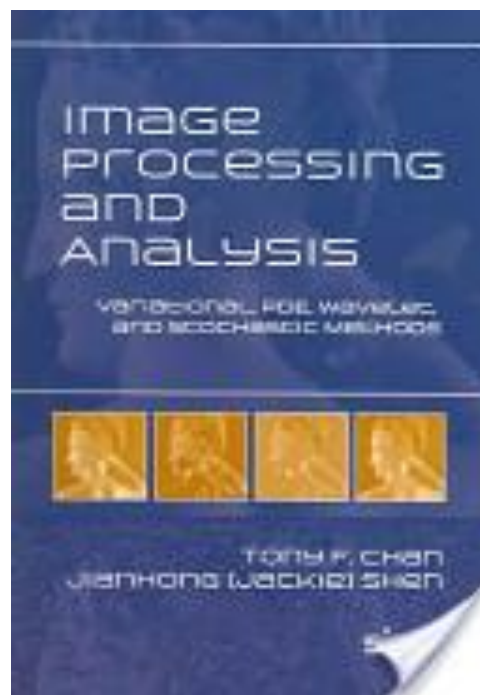




ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ ΤΕΙ ΗΛΕΚΤΡΟΝΙΚΗΣ ΧΑΝΙΩΝ

# Image Processing and Analysis

By Tony Chan, Jianhong Shen



**ΜΟΥΖΟΥΛΑΣ ΣΩΤΗΡΙΟΣ**

[sotism@windowslive.com](mailto:sotism@windowslive.com)

**ΣΚΕΥΟΣ ΚΟΚΚΙΝΙΔΗΣ**

[dj\\_exploder\\_3@hotmail.com](mailto:dj_exploder_3@hotmail.com)

**ΚΕΦΑΛΑΙΑ 2, 3, 5**

## ΚΕΦΑΛΑΙΟ 2

Παρόμοια με το προηγούμενο παράδειγμα (2.43) το οποίο είναι εμπνευσμένο από το νόμο περί κατανομής του Γκιμπς, ας πάρουμε το  $\varepsilon = h^\alpha / \beta$  για κάποια σταθερή παράμετρο  $\beta$  της εντολής 1. Στη συνέχεια στην ακόλουθη εντολή του  $h$  όπου  $h \rightarrow 0$

$$P_{ij,(i\pm 1,j)}^\varepsilon = \frac{1}{Z^\varepsilon} e^{-\beta|u_x|^\alpha}_{i\pm\frac{1}{2},j}, \quad P_{ij,(i,j\pm 1)}^\varepsilon = \frac{1}{Z^\varepsilon} e^{-\beta|u_y|^\alpha}_{i,j\pm\frac{1}{2}},$$

$$D_1(\nabla u) = \frac{h^2}{\Delta t Z^\varepsilon} e^{-\beta|u_x|^\alpha}, \quad D_2(\nabla u) = \frac{h^2}{\Delta t Z^\varepsilon} e^{-\beta|u_y|^\alpha},$$

με το  $\Delta t$  να υποδουκνύει τον χρόνο αναμονής ο οποίος σχετίζεται με τη μονοβηματική μετάβαση της διαδικασίας. Ας αφήσουμε το  $\Delta t$ ,  $h \rightarrow 0$  με τέτοιο τρόπο ώστε τόσο το  $D_1$  όσο και το  $D_2$  να συγκλίνουν. Έπειτα λαμβάνουμε τη μη γραμμική ανισότροπη διάχυση

$$u_t = \nabla \cdot D \nabla u = (D_1(\nabla u)u_x)_x + (D_2(\nabla u)u_y)_y,$$

Με το  $D = \text{diag}(D_1, D_2)$  να είναι ένα διαγώνιο κύκλωμα μήτρας. Περαιτέρω, εφόσον  $h \rightarrow 0$  η τμηματική συνάρτηση  $Z^\varepsilon$  συγκλίνει στο

$$Z = 1 + 2(e^{-\beta|u_x|^\alpha} + e^{-\beta|u_y|^\alpha}).$$

Συνεπώς αν αφήσουμε το  $h$ ,  $\Delta t \rightarrow 0$  ώστε το  $h^2 / \Delta t \rightarrow 3$ , τότε

$$D_1 = \frac{3e^{-\beta|u_x|^\alpha}}{1 + 2(e^{-\beta|u_x|^\alpha} + e^{-\beta|u_y|^\alpha})}, \quad D_2 = \frac{3e^{-\beta|u_y|^\alpha}}{1 + 2(e^{-\beta|u_x|^\alpha} + e^{-\beta|u_y|^\alpha})}. \quad (2.47)$$

Η τεχνητή επιλογή του αριθμού 3 οδηγεί στον απλό αλλά περιοριστικό υπολογισμό:  $D_1 < 1$  και  $D_2 < 1$ .

## 2.6 Κυμάτια και Ανάλυση Πολλαπλής διακριτικής ικανότητας

Σε αυτή την ενότητα θέλουμε να κάνουμε μια περιληπτική εισαγωγή στην κλασσική θεωρία των κυματιών [96, 204, 215, 290]. Οι πιο πρόσφατες προόδους στα γεωμετρικά και μη γραμμικά κυμάτια μπορούν να διαβαστούν π.χ. [41, 107, 108, 155, 248].

### 2.6.1 Αναζήτηση Νέων Εργαλίων Ανάλυσης Εικόνων

Σε αντίθεση με άλλα σήματα, οι εικόνες είναι αποτέλεσμα της τυχαίας τοποθέτησης ξεχωριστών αντικειμένων στον κόσμο του 3-D. Τα πιο έντονα χαρακτηριστικά αυτών των αντικειμένων είναι φανερά στην μεγάλη ποικιλία χωρικών κλιμάκων και τις ευκόλως αναγνωρίσιμες (στην ανθρώπινη

όραση) διαφορές όσον αφορά την γεωμετρική επιφάνεια, την ανακλαστικότητα και διάφορα μοτίβα υφών επιφάνειας.

Η ικανότητα το μη συσχετισμού και η αναγνώριση ξεχωριστών εικόνων από αντικείμενα καθώς και τα μοτίβα τους είναι επομένως απαραίτητη για την ανάλυση εικόνων και όρασης. Τα κυμάτια και η σχετικές τεχνικές έχουν ικανοποιήσει αυτές τις απαιτήσεις επιτυχώς μέχρι στιγμής.

Κυριολεκτώντας ένα ξεχωριστό κυμάτιο είναι ένα τοπικό μικρό κύμα. Λειτουργεί σαν ένας εικονικός νευρώνας ο οποίος αφήνεται έντονα κάθε φορά που του παρουσιάζονται τοπικά οπτικά χαρακτηριστικά. Λόγω της κατά τόπους προσαρμογής, μπορεί να απαντήσει έντονα μόνο όταν το παράθυρό του απαθανατίζει τα στοχευμένα χαρακτηριστικά στο εσωτερικό.

Η *ανάλυση κυματίων* μελετά τον σχεδιασμό, την οργάνωση και την ανάλυση τέτοιων κυματίων και επιτυγχάνει επαρκή υπολογιστικά σχήματα. Ένα σημαντικό μέρος του σκοπού της είναι να δημιουργήσει ένα επιστημονικό μοντέλο ανθρώπινης και μηχανικής όρασης και να εκτελέσει επιτυχώς διάφορα έργα επεξεργασίας εικόνας.

Για να κατανοήσουμε και να εκτιμήσουμε τα χαρακτηριστικά και τα πλεονεκτήματα των κυματίων καλύτερα, ας επιστρέψουμε για λίγο στην ανάλυση των συμβατικών σημάτων τα οποία δημιουργούνται από συνηθισμένα φυσικά συστήματα. Κοντά στην κατάσταση ισορροπίας τέτοια συστήματα μπορούν συνήθως να υπολογιστούν ορθά από γραμμικά συστήματα. Κατόπιν μια είσοδος  $f$  και μια έξοδος  $u$  συνδέονται με ένα γραμμικό σύστημα  $A$ :

$u = Af$  στην απουσία θορύβου.

Για κάθε σύστημα όπως η χορδή μιας κιθάρας, το φωνητικό κανάλι ενός ανθρώπου ή η επιφάνεια ενός ντραμ, το χωρικό μέρος του  $A$  μπορεί συχνά να παρουσιαστεί σωστά από ένα ελλειπτικό τελεστή δευτέρου βαθμού

$$L = \nabla \cdot D(x, t) \nabla + b(x, t)$$

Όπου το  $x$  και το  $t$  δηλώνουν τις χωρικές και τις χρονικές μεταβλητές και το  $D$  και το  $b$  διευκρινίζουν την φυσική κατάσταση του συστήματος στόχου. Στην πράξη είναι επίσης απαραίτητο να διευκρινιστούν κατάλληλες οριακές συνθήκες.

Αν η φύση του συστήματος είναι ανεξάρτητη από τον χρόνο  $t$ :  $D = D(x)$  και  $b = b(x)$  η ενδογενής του κατάσταση ή ιδιομεθόδοι είναι αναλλοίωτοι στο  $L$  και διευκρινίζονται από το πρόβλημα ιδιοτιμής

$L\phi_\lambda = \lambda\phi_\lambda$  με κατάλληλες οριακές συνθήκες,

όπου η ιδιοτιμή  $\lambda$  αντικατοπτρίζει το ενεργειακό επίπεδο της συσχετιζόμενης κατάστασης. Ως προς το γένος, το σύστημα κατέχει μια συλλογή από ιδιομεθόδους,

$(\phi_n, \lambda_n)$ , οι οποίες διατάζονται από το  $\lambda_0 \leq \lambda_1 \leq \dots$ .

Η γενική θεωρία των γραμμικών ελλειπτικών τελεστών δευτέρου βαθμού [117, 132] υποστηρίζει ότι  $\lambda_n \rightarrow \infty$  καθώς το  $n \rightarrow \infty$  και το  $(\phi_n)_n$  δίνουν μια φυσική γραμμική βάση για τη μελέτη της δημιουργίας και της εξέλιξης σημάτων υπό το γραμμικό σύστημα  $A$ , το οποίο είναι συνήθως συνδεδεμένο με το  $\partial_t - L$  αν το  $A$  είναι πλήρως διαχυνόμενο ή  $\partial_{tt} - L$  αν το  $A$  είναι πλήρως μεταφορικό.

Για παράδειγμα ας σκεφτούμε τη μικρή μήκους δόνηση μιας χορδής με μήκος  $a$  και τα άκρα δεμένα. Οι ιδιομεθόδοι του δίνονται από το

$$-u_{xx} = \lambda u, \quad u(0) = u(a) = 0.$$

Έπειτα είναι ευρέως γνωστό ότι

$$\lambda_n = \left(\frac{\pi}{a}\right)^2 n^2, \quad n = 1, 2, \dots \quad \text{and} \quad \phi_n = \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right), \quad (2.48)$$

έτσι

ώστε το  $(\phi_n)_n$  να είναι μια ορθογωνική βάση αρμονικών κυμάτων για το  $L^2(0, a)$ . Όπως είναι εμφανές από τις εξισώσεις, αυτές οι αρμονικές μέθοδοι συνδυάζουν άψογα τόσο τον νόμο της κίνησης (των χορδών) όσο και τη συνολική γεωμετρία (π.χ. μέγεθος  $a$ ), και για αυτό δεν μπορούν να είναι τοπικές.

Η μη τοπικότητα μπορεί έπειτα να γίνει εμφανής θεωρώντας το πρόβλημα διάχυσης σήματος

$$u_t = u_{xx}, \quad 0 < x < a, \quad u(0) = u(a) = 0,$$

με την αρχική μορφή να δίδεται από το δέλτα του Ντάϊρακ  $u(x, 0) = f = \delta(x - a/2)$ , μια εξιδανικευμένη τοπική παρόρμηση στο μέσο σημείο. Η σηματική διάχυση μπορεί να εκφραστεί. Προσέξτε ότι η αρχική παρόρμηση μολύνει εξίσου όλες τις περιπτώσεις ιδιομεθόδους καθώς

$$a_{2n+1}(0) = \langle \phi_{2n+1}, \delta(x - a/2) \rangle = \pm \sqrt{\frac{2}{a}}.$$

Συνεπώς, η χωρική τοπικότητα δεν είναι ενδογενής για τέτοια συστήματα. Για να επιτευχθεί αυτό πρέπει να στραφούμε στα κυμάτια για περιορισμένη σηματική ανάλυση ή επεξεργασία.

## 2.6.2 Η Αρχική Θεωρία των Άκρων και τα κυμάτια του Μαρς

Η άκρη είναι το πιο συνηθισμένο και σημαντικό οπτικό χαρακτηριστικό στις εικόνες και καθορίζει και διαχωρίζει τα αντικείμενα και αφήνει σοβαρά στοιχεία για τις 3-D χωρικές εντολές [234]. Συνεπώς ο σωστός καθορισμός και η εξαγωγή των άκρων από εικόνες 2-D όπως αναπτύχθηκε πρώτα από τον Marr [210] και τους Marr και Hildreth [211] και έπειτα αναπτύχθηκε από πολλούς άλλους [44, 130, 165, 226] είναι ένα από τα πιο ουσιαστικά προβλήματα στην ανάλυση εικόνων και όρασης.

Ο Μαρ όρισε τις άκρες μιας εικόνας  $u$  ως τα σημεία όπου  $\Delta u_\sigma = 0$ , με τις μηδενικές διασταυρώσεις της λαπλασιανής της εικόνας  $u$  να σβήνουν στη κλίμακα  $\sigma$ . Πιο συγκεκριμένα, έστω ότι  $x = (x_1, x_2)$  και  $r = |x|$ , και έστω

$$G(x) = \frac{1}{2\pi} e^{-\frac{x^2}{2}} \quad \text{and} \quad G_\sigma(x) = \frac{1}{\sigma^2} G\left(\frac{x}{\sigma}\right) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, \quad (2.49)$$

ώστε το

$G_\sigma$  είναι μια κάθετα συμμετρική πιθανότητα πυκνότητα με διακύμανση  $2\sigma^2$ . Έπειτα  $u_\sigma = u * G_\sigma$ , και

$$\Delta u_\sigma = u * \Delta G_\sigma = u * \psi_\sigma,$$

όπου το  $\psi_\sigma = \Delta G_\sigma$  αποκαλείται τώρα συχνά η θεωρία κυματίων του Μαρ:

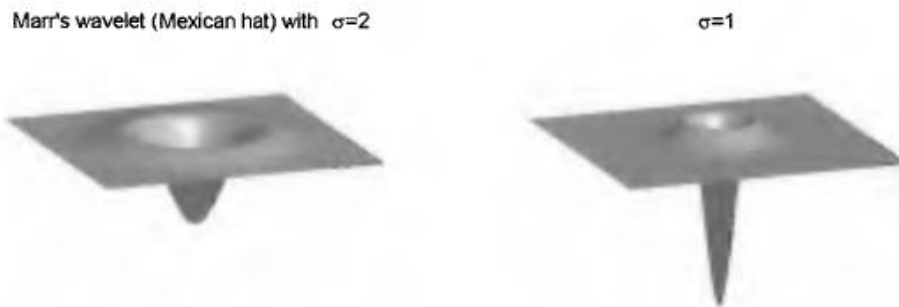
$$\psi_\sigma = \frac{1}{\pi\sigma^4} \left( \frac{x^2}{2\sigma^2} - 1 \right) e^{-\frac{x^2}{2\sigma^2}} = \frac{1}{\sigma^4} \psi_1\left(\frac{x}{\sigma}\right). \quad (2.50)$$

Σημειώσ

τε ότι το  $\psi_\sigma$  ικανοποιεί την κατάσταση διέλευσης υψηλών ταχυτήτων  $\int_{R^2} \psi_\sigma = 0$ . Στην Εικόνα 2.14 βρίσκονται σχεδιασμένα δυο παραδείγματα από τα κυμάτια του Μαρ με διαφορετικά  $\sigma$ .

Η ιδέα πίσω από τη θεωρία μηδενικής διασταύρωσης άκρων του Μαρ έχει ως ακολούθως. Ας υποθέσουμε ότι το  $u$  έχει μια ιδανική άκρη κατά βαθμίδες κατά μήκος ενός ευθέους τμήματος γραμμής  $l$ . Έστω ότι το  $\mathbf{t}$  συμβολίζει την εφαπτόμενη κατεύθυνση μονάδας του  $l$  και το  $\mathbf{n}$  την κανονική του μονάδα.

$$K(x, y)$$



**Εικόνα 2.14** Δυο παραδείγματα κυματιδίων του Μαρ («Μεξικάνικα καπέλα») όπως στο 2.50.

Έπειτα το  $(\mathbf{t}, \mathbf{n})$  μας δίνει τοπικές καρτεσιανές συντεταγμένες και

$$\Delta = \frac{\partial^2}{\partial \mathbf{t}^2} + \frac{\partial^2}{\partial \mathbf{n}^2},$$

λόγω της περιστροφικής αμεταβλητότητας της λαπλασιανής. Ας υποθέσουμε ότι το μήκος του  $\sigma \ll l$ . Τότε το  $u_\sigma$  εξακολουθεί να είναι μεταβατικά αμετάβλητο προς την κατεύθυνση  $\mathbf{t}$  (μέχρι μια εκθετικά μικρή απόκλιση). Συνεπώς έχουμε

$$\frac{\partial u}{\partial \mathbf{t}} \simeq 0, \quad \frac{\partial^2 u}{\partial \mathbf{t}^2} \simeq 0, \quad \text{and} \quad \Delta u_\sigma \simeq \frac{\partial^2 u_\sigma}{\partial \mathbf{n}^2}.$$

Ως εκ τούτου οι μηδενικές διασταυρώσεις του  $\Delta u_\sigma$  είναι τα σημεία κλίσης προς την κατεύθυνση  $\mathbf{n}$  όπου το  $|\nabla u_\sigma| = |\partial u_\sigma / \partial \mathbf{n}|$  φτάνει στο μέγιστο σημείο. Για να γίνει διαχωρισμός ανάμεσα στα σημαντικά και ασήμαντα σημεία άκρης στις αργά μεταβαλλόμενες περιοχές στις γενικές εικόνες, φαίνεται απαραίτητη η μελέτη και της πληροφορίας κλίσης  $\nabla u_\sigma$  [44].

Σε αυτό το πολύ γνωστό παράδειγμα, τα κυμάτια Μαρ  $\psi_\sigma$  χρησιμοποιούνται απλά σαν τοπικά εργαλεία ανίχνευσης. Εδώ ο όρος «κυμάτιο» χρησιμοποιείται με την ευρύτερη έννοια των Κρόνλαντ-Μάρτινιτς, Μορλέτ και Γκρόσμαν [185], δηλαδή οποιαδήποτε λειτουργία  $\psi(x)$  που φθίνει ικανοποιητικά γρήγορα στο  $\infty$  και ικανοποιεί την κατάσταση υψηλών συχνοτήτων

$$\int_{R^2} \psi = 0.$$

### 2.6.3 Παραθυρωμένη Ανάλυση Συχνοτήτων και τα κυμάτια του Γκάμππορ.

Σε σύγκριση με τη συμβατική ανάλυση συχνοτήτων (Φουριέ) η παραθυρωμένη ανάλυση συχνοτήτων προωθεί την τοπικότητα εφαρμόζοντας τον μετασχηματισμό του Φουριέ στο παραθυρωμένο σήμα τις εικόνας  $w(x - a)u(x)$ , όπου το  $w(x) = w(x_1, x_2)$  αντιπροσωπεύει ένα τοπικό παράθυρο και το  $a = (a_1, a_2)$  το κινούμενο του κέντρο. Ένα παράθυρο  $w(x)$  είναι συνήθως μια ραγδαίως φθίνουσα πραγματική λειτουργία με  $u(x) = u(x_1, x_2)$  Ο παραθυρωμένος μετασχηματισμός του Φουριέ (πΜΦ) μιας δεδομένης εικόνας  $u(x) = u(x_1, x_2)$  καθορίζεται ως

$$U(a, k) = \int_{\mathbb{R}^2} u(x)w(x - a)e^{-i\langle k, x \rangle} dx,$$

το οποίο είναι μια τετραδιάστατη (4-D) λειτουργία της θέσης  $a = (a_1, a_2)$  και της χωρικής συχνότητας  $k = (k_1, k_2)$ . Θα μπορούσε να γίνει περίπου κατανοητός ως η δύναμη της χωρικής συχνότητας  $k$  τοπικά στο  $x = a$ .

Ο πΜΦ (2.51) είναι επίσης η προβολή τις εικόνας  $u$  πάνω σε μια συλλογή τοπικών κυμάτων

$$\psi_{a,k}(x) = w(x - a)e^{i\langle k, x \rangle}, \quad a, k \in \mathbb{R}^2,$$

με μετακινούμενο κέντρο  $a$  και τοπική συχνότητα  $k$ . Συνεπώς, τα  $\psi_{a,k}$  αποκαλούνται κανονικά κυμάτια. Όταν το  $w$  είναι μια βαθμονομημένη γκαουσιανή, είναι ευρέως γνωστά ως τα κυμάτια Γκάμππορ.

Η διαδικασία αναδόμησης διεξάγεται ως αν τα  $(\psi_{a,k})_{a,k}$  ήταν μια ορθογώνια βάση:

$$u(x) = \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} U(a, k)\psi_{a,k}(x) \frac{dk da}{(2\pi)^2},$$

κυρίως χάρη στο γεγονός ότι από το (2.51) ο αντίστροφος μετασχηματισμός του Φουριέ στο  $x$  μόνο μπορεί να επαναφέρει το  $w(x - a)u(x)$  και ότι

$$\int_{\mathbb{R}^2} w^2(x - a)da = \int_{\mathbb{R}^2} w^2 = 1.$$

Όσον αφορά τα κυμάτια του Μαρ, κάποιος θα μπορούσε να εισάγει πολλαπλές βαθμίδες στον πΜΦ επαναβαθμίζοντας το δεδομένο παράθυρο. Ορίστε

$$w_\sigma(x_1, x_2) = \frac{1}{\sigma} w\left(\frac{x}{\sigma}\right)$$

και τα  $\psi_{a,k,\sigma}(x) = w_\sigma(x-a)e^{i(k \cdot x)}$  τα οποία οδηγούν στα πιο γενικά κυμάτια του Γκάμπορ. Έπειτα η πολυβαθμίδα πΜΦ καθορίζεται ως

$$U_\sigma(a, k) = \langle u(x), \psi_{a,k,\sigma}(x) \rangle.$$

Όπως και προηγουμένως το  $u(x)$  θα μπορούσε να αναδομηθεί τέλεια από το  $U_\sigma$  σε οποιαδήποτε ενιαία κλίμακα.

Το σοβαρότερο μειονέκτημα αυτών των κυματιών του Γκάμπορ είναι ότι η κλίμακα  $\sigma$  και η χωρική συχνότητα  $k$  συνδυάζονται με έναν κάπως άτεχνο τρόπο (π.χ. μέσω μη συσχετισμένου τανυστικού προϊόντος). Στην πράξη θα έπρεπε να είναι συζευγμένα: οι υψηλές συχνότητες είναι πιο απαραίτητες για την επίλυση χαρακτηριστικών μικρών βαθμίδων ενώ οι χαμηλές συχνότητες απο μόνες του αρκούν για χαρακτηριστικά μεγάλων βαθμίδων.

Ένα άλλο θέμα αφορά στον πλεονασμό του βήματος ανάλυσης (2.51) και τη σταθερότητα στο βήμα σύνθεσης (2.53). Στην τετραδιάστατη παράμετρο διαστήματος-συχνότητας το διάστημα  $(a, k) \in \mathbb{R}^2 \times \mathbb{R}^2$  φαίνεται προφανώς πλεονάζων στην κρυπτογράφιση εικόνων 2-D. Επιπλέον, η έλλειψη ορθογωνιότητας (ή σχεδόν ορθογωνιότητας) θα μπορούσε να προκαλέσει πιθανώς σοβαρά προβλήματα σταθερότητας λόγω των θορύβων εικόνας ή των σφαλμάτων προσέγγισης.

Αυτές οι μελέτες έχουν εμπνεύσει περαιτέρω βελτίωση στη θεωρία των κυματιών.

#### 2.6.4 Σύζευξη Συχνότητας-Παραθύρου: Τα Κυμάτια Μαλβάρ-Γουίλσον

Για χάρη της απλότητας θα δουλέψουμε μόνο με μονοδιάστατα σήματα σε όλο τον πραγματικό άξονα  $\mathbb{R}$ .

Ας υποθέσουμε ότι έχουμε μια συλλογή παραθύρων

$$w_n(x), \quad n = 0, \pm 1, \pm 2, \dots$$

που καλύπτει ολόκληρο το διάστημα έτσι ώστε η στήριξη  $w_n = [a_n, b_n]$  να είναι πεπερασμένα διαστήματα με μεταβλητά μήκη  $l_n = b_n - a_n$ .

Έπειτα κάθε παραθυρωμένη εικόνα  $u_n(x) = w_n(x)u(x)$  διαμορφώνονται σε  $[a_n, b_n]$ . Με την περιοδική επέκταση  $l_n$ ,  $u_n(x) \rightarrow \hat{u}_n(x), \hat{u}_n$

μπορεί να επανακτηθεί πλήρως (στο  $L^2(a_n, b_n)$ ) από τους συντελεστές του Φουριέ:

$$U(n, k) = \langle \hat{u}_n(x), e^{i \frac{2\pi}{l_n} kx} \rangle = \langle u(x), \psi_{n,k}(x) \rangle, \quad (2.54)$$

όπου το

$$\psi_{n,k}(x) = w_n(x)e^{i \frac{2\pi}{l_n} kx}$$

εσωτερικό προϊόν βρίσκεται στο  $L^2(a_n, b_n)$  και τα



αποτελούν καινούρια κυμάτια. Προσέξτε τη σύζευξη μεταξύ μεγεθών παραθύρων  $l_n$  και τις συνιστώσες συχνότητας.

$$\sum_n w_n^2(x) = 1 \text{ for all } x \in \mathbb{R}.$$

Ας υποθέσουμε όπως και προηγουμένως ότι

Τότε

$$\begin{aligned} \sum_{n \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \frac{1}{l_n} U(n, k) \psi_{n,k}(x) &= \sum_{n \in \mathbb{Z}} w_n(x) \sum_{k \in \mathbb{Z}} \frac{1}{l_n} \langle \hat{u}_n(y), e^{i \frac{2\pi}{l_n} k y} \rangle e^{i \frac{2\pi}{l_n} k x} \\ &= \sum_{n \in \mathbb{Z}} w_n(x) \hat{u}_n(x) = \sum_{n \in \mathbb{Z}} u_n(x) w_n(x) \\ &= u(x) \sum_{n \in \mathbb{Z}} w_n^2(x) = u(x). \end{aligned}$$

Δηλαδή

, η τέλεια αναβάθμιση επιτυγχάνεται και πάλι από τους συντελεστές κυματίων  $U(n, k)$ ,  $n, k \in \mathbb{Z}$ . Σε σύγκριση με το συνεχόμενο πΜΦ  $(\alpha, k)$ ,  $\alpha, k \in \mathbb{Z}$ .

που αναλύσαμε στην προηγούμενη ενότητα ο πλεονασμός μειώθηκε σημαντικά αλλά δεν αφαιρέθηκε εντελώς. Ακόμη πιο σημαντικό, στη νέα ομάδα κυματίων  $\psi_{n,k}(x)$  η χωρική ανάλυση έγινε πιο επαρκής μετά την ενσωμάτωση των μεγεθών παραθύρων  $l_n$  στις υψηλές συχνότητες.

Πώς θα μπορούσε ο πλεονασμός να παραμείνει σε αυτή τη βελτιωμένη εκδοχή του ξεχωριστού πΜΦ; Βρίσκεται ριζωμένο στην αναπόφευκτη υπερκάλυψη των παραθύρων όπως αυτό απαιτείται από τη συνθήκη κανονικοποίησης

$$\sum_{n \in \mathbb{Z}} w_n^2(x) = 1, \quad x \in \mathbb{R},$$

και τις απαραίτητες συνθήκες ομαλότητας που επιβάλλονται στα παράθυρα για ένα αριθμό αναλυτικών καθώς και υπολογιστικών λόγων.

Στην περίπτωση που τα παράθυρα  $(w_n)_n$  είναι «σκληροί κόφτες»,

$$w_n(x) = 1_{[a_n, a_{n+1})}(x), \quad \dots < a_n < a_{n+1} < \dots, \quad (2.55)$$

π.χ.

δείκτης μη-ομαλών λειτουργιών που σχετίζονται με ένα διαχωρισμό του όλου διαστήματος, κάποιος θα μπορούσε εύκολα να δείξει ότι ο πλεονασμός έχει όντως εξαφανιστεί εντελώς εφόσον το

$(\psi_{n,k} / \sqrt{l_n})_{n,k}$  είναι μια ορθογωνική βάση του  $L^2(\mathbb{R})$ . Αλλά μια τόσο φτηνή ορθογωνικότητα έχει το βαρύ κόστος της κατά προσέγγιση ακριβείας: για μια γονική ομαλή εικόνα  $u(x)$ , μια απλή μη συνεχόμενη γραμμική λειτουργία για παράδειγμα, το σφάλμα προσέγγισης φθίνει με αργό ρυθμό του  $O(|k|^{-1})$ . Συνεπώς, πολύ πιο «απαλά» ή ομαλά παράθυρα προτιμώνται για να ασφαλίσουν υψηλή ακρίβεια στην ανάλυση και τη σύνθεση εικόνων.

Τα ερωτήματα κλειδί είναι τα εξής: Μπορεί κάποιος να σχεδιάσει όντως μια ορθογωνική βάση που είναι πολύ όμοια στο  $(\psi_{n,k})_{n,k}$  με τη χρήση ομαλών παραθύρων; και από ποιες απόψεις μπορεί τα  $\psi_{n,k}$  να σπλιθωθούν περαιτέρω; Η απάντηση όσον αφορά τον τομέα της ευστάθειας είναι θετική. Ανάμεσα σε πολλά από τα ήδη υπάρχοντα σχέδια, κάνουμε μια σύντομη εισαγωγή στην αξιοσημείωτη προσέγγιση του Μαλβάρ [205] και Γουίλσον [321] η οποία οδηγεί σε αυτό που αποκαλείται τα κυμάτια του Μάλβαρ – Γουίλσον και οι παραλλαγές τους [91, 215].

Πρώτα επιλέγουμε ένα περίγραμμα ομαλού παραθύρου  $w(x)$  που ικανοποιεί (βλέπε εικόνα 2.15):

1. το  $w$  στηρίζεται στο  $[-\pi, 3\pi]$  και είναι συμμετρικό γύρω από το  $x = \pi$ :  $w(2\pi - x) = w(x)$

2. απαγορεύεται στο  $[-\pi, \pi]$ ,  $w(x)$  ικανοποιεί τη συνθήκη μη-οδόντωσης  $w^2(x) + w^2(-x) = 1$ .

Δημιουργείται μια σειρά παραθύρων από μεταφράσεις του τύπου  $2\pi$ :  $w_n(x) = w(x - 2n\pi)$ ,  $n = 0, \pm 1, \pm 2, \dots$ . Τότε η συμμετρία και η μη-οδοντική συνθήκη οδηγούν στο

$$\sum_n w_n^2(x) = 1, \quad x \in \mathbb{R}.$$

Προσέξτε ότι το άπειρο άθροισμα αποτελείται μόνο από το πολύ δυο όρους εφόσον η στήριξη  $w_n$  και η στήριξη  $w_m$  δεν υπερκαλύπτονται ποτέ για το  $|n - m| > 1$ .

Αν ακολουθείτο η παλιά δομή (2.54) κάποιος θα έπρεπε να χρησιμοποιήσει  $e^{i\frac{k}{2}x}$ ,  $k \in \mathbb{Z}$ , για να συσχετιστεί με κάθε παράθυρο  $w_n(x)$  αφού  $l_n = l_0 = 2\pi$  ή ισοδύναμα,

$$\left\{ \cos \frac{k}{2}x : k = 0, 1, \dots \right\} \text{ and } \left\{ \sin \frac{k}{2}x : k = 1, 2, \dots \right\}. \quad (2.56)$$

Αλλά

δύο παρακαίμενα παράθυρα  $w_n$  και  $w_{n+1}$  υπερκαλύπτονται ξεκάθαρα και αναπόφευκτα εισάγουν τον πλεονασμό που μπορεί, παράλα αυτά, να αφαιρεθεί έξυπνα από μια *εναλλασσόμενη* «υποδειγματοληψία» του συντελεστή 2. Με άλλα λόγια, συσχετισμένο με οποιοδήποτε άρτιο παράθυρο  $w_{2n}(x)$  κάποιος εφαρμόζει μόνο το μισό από το (2.56),

$$\left\{ \cos \frac{k}{2}x : k = 0, 1, \dots \right\},$$

ενώ για οποιοδήποτε περιττό παράθυρο  $w_{2n+1}$  κάποιος εφαρμόζει το άλλο μισό,

$$\left\{ \sin \frac{k}{2}x : k = 1, 2, \dots \right\}.$$

Συνεπώς, έχουμε αντλήσει την ομάδα των λεγόμενων κυμάτων Μάλβαρ - Γουίλσον:

$$\psi_{2n,k} = c_k w_{2n}(x) \cos \frac{k}{2}x, \quad k \geq 0, \quad \text{and} \quad \psi_{2n+1,k} = c_k w_{2n+1}(x) \sin \frac{k}{2}x, \quad (2.57)$$

όπου  $c_0$

= 1 και  $c_k = \sqrt{2}, k \geq 1$  για την κανονικοποίηση  $L^2$ .

Το ωραίο είναι ότι τα κυμάτια του Μάλβαρ – Γουίλσον  $(\psi_{n,k})_{n,k}$  αποτελούν μια ορθογωνική βάση στο  $L^2(R)$ ! Η διαορθογωνιότητα ανάμεσα σε δυο παράλληλα κυμάτια

$$\langle \psi_{n,k}(x), \psi_{n+1,l}(x) \rangle = 0$$

καταλήγει από τη συμμετρία (ή την ομαλότητα) του πίνακα  $u(x)$  και την περιπτότητα του προϊόντος ενός συνημίτονου με ένα ημίτονο, καθώς η ενδοορθογωνιότητα ανάμεσα σε δύο κυμάτια μέσα σε ένα κοινό παράθυρο

$$\langle \psi_{n,k}(x), \psi_{n,l}(x) \rangle = 0$$

είναι εγγυημένη από τη συνθήκη μη-οδόντωσης του πίνακα καθώς και από το γεγονός ότι τόσο

$$\left\{ c_k \cos \frac{k}{2}x : k = 0, 1, \dots \right\} \quad \text{and} \quad \left\{ c_k \sin \frac{k}{2}x : k = 1, 2, \dots \right\}$$

είναι

ορθογωνικές βάσεις για το  $L^2[0, 2\pi]$ . Η τελευταία δήλωση πηγάζει από το γεγονός ότι αυτά τα συνημίτονα είναι ακριβώς οι ιδιομεθόδοι του συστήματος Στερμ – Λιούβιλ

$$-u_{xx} = \lambda u, \quad 0 \leq x \leq 2\pi, \quad u'(0) = u'(2\pi) = 0,$$

ενώ τα

ημίτονα είναι οι ιδιομεθόδοι του

$$-u_{xx} = \lambda u, \quad 0 \leq x \leq 2\pi, \quad u(0) = u(2\pi) = 0.$$

Αυτή η εναλλασσόμενη σύνδεση συνημίτονου – ημίτονου προτάθηκε για πρώτη φορά από τον νομπελίστα Κένεθ Γουίλσον στη μελέτη της θεωρίας επανακανονικοποίησης ομάδας [321], αργότερα αναπτύχθηκε συστηματικά στο αξιοσημείωτο άρθρο των Ντόμπρετς, Τζαφάρτ και Ζουρνέ [97] και επανακαλύφθηκε ανεξάρτητα στην επεξεργασία ψηφιακού σήματος από τον Μάλβαρ [205].

Τα κυμάτια Μάλβαρ-Γουίλσον θα μπορούσαν να επεκταθούν σε μια ευρέα τάξη ορθογωνικών κυμάτων διαστήματος – συχνότητας όπως έγινε από τους Κόιφμαν και Μέγιερ [91]. Για παράδειγμα, κάποιος θα μπορούσε να αντικαταστήσει την αμετάβλητη μετάφραση των παραθύρων

$w_n(x) = w(x - na)$  για ορισμένα είναι κατάλληλο το  $a$

μέσω αμετάβλητης διαστολής

$w_n(x) = w(\lambda^n x)$  για ορισμένα  $\lambda > 0, \lambda \neq 2$  για παράδειγμα.

Τέτοια παράθυρα έχουν μεταβλητό μήκος παραθύρων και ανάλογα τα συνδεδεμένα κύματα συνημίτονου και ημίτονου θα έπρεπε να ιεραρχηθούν αναλόγως. Πολύ πιο γενικά, κάποιος θα

μπορούσε να εγκαταλείψει εντελώς ένα πίνακα μονού παραθύρου και την μεταφραστικά ή ιεραρχικά αναλλοίωτη ιδιοκτησία και να σχεδιάσει ανεξάρτητα κάθε παράθυρο  $w_n$ . Με τη προϋπόθεση ότι κάποιος

1. διατηρεί τη συνθήκη μη οδόντωσης για ενδοορθογωνιότητα και
2. συντονίζει καλά τη συμπεριφορά συμμετρίας οποιουδήποτε ζευγαριού γειτονικών παραθύρων και των αρμονικά συνδεδεμένων κυματίων στην επικαλυπτόμενη περιοχή για ενδοορθογωνιότητα

μια ορθογωνική βάση κυματίων διαστήματος – συχνότητας μπορεί να επιτευχθεί εύκολα με αυτό το σχέδιο.

### 2.6.5 Το πλαίσιο Ανάλυσης Πολλαπλής ευκρίνειας Εικόνας (MRA)

Από τους σχεδιασμούς του Γκείμπορ μέχρι αυτούς του Μάλβαρ – Γουίλσον, τα κυμάτια ζουν ακόμη στην επιβλιτική σκιά της ανάλυσης συχνότητας του Φουριέ. Ιστορικά, η κατάσταση δεν παρουσίασε ουσιαστική στροφή μέχρι τη στιγμή που η Μέγιερ και Μαλάτ εισήγαγαν το ανεξάρτητο και γενικό πλαίσιο της ανάλυσης πολλαπλής ευκρίνειας εικόνας (MRA). Παρόλο που αρχικά πηγάζει από μεθοδολογίες επεξεργασίας σημάτων και εικόνας, το MRA δεν φαίνεται εντελώς ξένο στους επιστήμονες που εργάζονται σε φαινόμενα πολλαπλών κλιμάκων όπως τυρβώδεις ροές πολλαπλής κλίμακας και μέθοδοι πολλαπλών πλεγμάτων.

Οι συμβολές του πλαισίου MRA στην ανάπτυξη κυματιδίων υπήρξαν αξιοσημείωτες για πολλούς σημαντικούς λόγους. Κατά πρώτον, μερικές εφευρετικές νέες ιδέες στη σύγχρονη επεξεργασία σημάτων και εικόνας είναι υφασμένες στις βασικές αρχές του MRA. Κατά δεύτερο το MRA δίνει μια συστηματική προσέγγιση για τη δόμηση και την ανάλυση κυματιδίων. Κυρίως, επιτρέπει γρήγορες και αποτελεσματικές εφαρμογές για την ανάλυση και σύνθεση εικόνων από κυματίδια. Τέλος και πάνω από όλα βγάζει την θεωρία των κυματιδίων από τη σκιά της ανάλυσης συχνότητας του Φουριέ και τη βάζει σε ένα σταθερό και αυτάρκες θεμέλιο της ανάλυσης πολλαπλής κλίμακας.

Το MRA είναι ένα κομψό μαθηματικό πλαίσιο για την αντιπροσώπευση και ανάλυση εικόνων και γενικών σημάτων σε πολλαπλές κλίμακες. Ένα ορθογωνικό MRA του  $L^2(R)$  είναι μια διατεταγμένη αλυσίδα κλειστών υποδιαστημάτων:

$$\cdots \subseteq V_{-1} \subseteq V_0 \subseteq V_1 \subseteq \cdots, \quad (2.58) \quad \text{το οποίο}$$

ικανοποιεί τις τρεις ακόλουθες προϋποθέσεις:

1. [Πληρότητα]

$$\overline{\lim_{j \rightarrow \infty} V_j} = L^2(R), \text{ and } \lim_{j \rightarrow -\infty} V_j = \{0\}. \quad [\text{Διαδική}]$$

$$u(x) \in V_j \Leftrightarrow u(2x) \in V_{j+1}.$$

Ομοιότητα]

2. [Μετάφραση Πυρήνα] Υπάρχει μια λειτουργία (σελ. 81), έτσι ώστε το  $(\varphi(x-k))_k$  να είναι μια ορθογωνική βάση του  $V_0$ .

Συνεπώς αν ένα τέτοιο MRA υπάρχει πραγματικά, ολόκληρος ο κόσμος του  $L^2(R)$  μπορεί να κωδικοποιηθεί από μια μοναδική λειτουργία  $\varphi$  του «πυρήνα». Κατά ακρίβεια, ορίστε

$$\phi_{j,k} = 2^{j/2} \phi(2^j x - k), \quad j, k \in \mathbb{Z}.$$

Με βάση τη δυαδική ομοιότητα, το  $(\phi_{j,k})_k$  είναι επίσης μια ορθογωνική βάση του  $V_j$ . Από την τοποθέτηση της πληρότητας, οποιαδήποτε εικόνα  $u(x) \in L^2(\mathbb{R})$  μπορεί να υπολογιστεί σωστά σε οποιαδήποτε προθητή ακρίβεια από την προβολή του  $u_j = P_j u$  στο  $V_j$ :

$$u_j = P_j u = \sum_k \langle u, \phi_{j,k} \rangle \phi_{j,k}.$$

Με αυτή την έννοια στα αρχικά στάδια της θεωρίας των κυματιδίων, το  $\phi$  αποκαλείται συχνά το «πατρικό» κυματίδιο. Τώρα αποκαλείται συχνά η λειτουργία κλιμάκωσης, ή η λειτουργία διάταξης από κάποιους επιστήμονες υπολογιστικής.

Σαν μια απαραίτητη προϋπόθεση, η λειτουργία κλιμάκωσης πρέπει να ικανοποιεί την προϋπόθεση δυαδικής ομοιότητας

$$\phi(x) = 2 \sum_k h_k \phi(2x - k) = \sum_k (\sqrt{2} h_k) \phi_{1,k} \quad (2.59)$$

με τους συντελεστές  $(h_k)_k$  που δίνονται από

$$h_k = \frac{1}{\sqrt{2}} \langle \phi, \phi_{1,k} \rangle, \quad k \in \mathbb{Z}.$$

Στην πράξη, (2.59) μας δίνει την εξίσωση για το σχεδιασμό της λειτουργίας κλιμάκωσης  $\phi(x)$  από κατάλληλους συντελεστές  $(h_k)_k$  και αποκαλείται συχνά η δυαδική σχέση, τελειοποιημένη εξίσωση ή διεσταλμένη εξίσωση.

Μια άλλη χαρακτηριστική κατάσταση της λειτουργίας κλιμάκωσης είναι αυτή του

$\int_{\mathbb{R}} \phi \neq 0$ , υποθέτοντας ότι το  $\phi \in L^2(\mathbb{R}) \cap L^1(\mathbb{R})$ . Για να αποφευχθεί η άσκοπη τεχνικότητα, ας υποθέσουμε για λόγους απλότητας ότι το  $\phi$  υποστηρίζεται με συμπαγές τρόπο. (Γενικά ένας μέτριος ρυθμός πτώσης όπως το  $\phi = O(|x|^{-1-\epsilon})$  στο  $x = \pm\infty$ , συνδυασμένος με εργαλεία όπως το θεώρημα του Λεμπέσκ το οποίο κυριαρχείται από σύγκλιση θα εξακολουθεί να είναι ικανός να πετύχει την ακόλουθη απόδειξη.) Ας υποθέσουμε

διαφορετικά  $\int_{\mathbb{R}} \phi = 0$ . Τότε

$$\int_R \phi_{j,k} = 0, \quad j, k \in Z.$$

$$\int_R u = 1.$$

Ας πάρουμε οποιαδήποτε εικόνα  $u \in L^2$  η οποία υποστηρίζεται συμπαγώς και το

Τότε σε οποιαδήποτε δοσμένο επίπεδο  $j$  με  $u_j = P_j u$  της κλίμακας,

$$\int_R u_j = \sum_k \langle u, \phi_{j,k} \rangle \int_R \phi_{j,k} = 0, \quad (2.60)$$

εφόσον το άπειρο άθροισμα είναι στην πραγματικότητα πεπερασμένο. Από την άλλη, από την τοποθέτηση της πληρότητας του MRA,  $u_j \rightarrow u$  στο  $L^2$  σαν  $j \rightarrow +\infty$ . Μιας και όλα τα  $(u_j)_{j \geq 0}$  μπορούν εύκολα να αποδειχθούν ότι είναι ομοιόμορφα και συμπαγώς στηριγμένα κάποιος πρέπει να έχει σαν  $j \rightarrow \infty$ ,

$$u_j \rightarrow u \text{ in } L^1 \text{ and } \int_R u_j \rightarrow \int_R u = 1,$$

Τ

ο οποίο έρχεται σε αντίθεση με το (2.60).

Αυτή η φαινομενικά ασήμαντη κατάσταση έχει θεμελιώδες επιδράσεις. Κατά πρώτον επιτρέπει την κανονικοποίηση της λειτουργίας κλιμάκωσης με το  $\int_R \phi = 1$ . Έπειτα, η εξίσωση διαστολής

$$1 = \sum_k h_k,$$

(2.59) που αποκαλείται η συνθήκη βαθυπερατού στην επεξεργασία ψηφιακού σήματος όταν  $h = (h_k)_k$  θεωρείται ψηφιακό φίλτρο. Κατά δεύτερον, η εφαρμογή του μετασχηματισμού του Φουριέ στην εξίσωση διαστολής (2.59) οδηγεί στο

$$\hat{\phi}(\omega) = H\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right),$$

όπου  $H(\omega) = \sum_k h_k e^{-ik\omega}$  είναι η παλμική απόκριση του φίλτρου  $h$ . Σε συνδυασμό με  $\hat{\phi}(0) = 1$ , η επανάληψη του (2.61) δίνει τη ρητή έκφραση της λειτουργίας κλιμάκωσης στον τομέα συχνότητας:

$$\hat{\phi}(\omega) = H\left(\frac{\omega}{2}\right) H\left(\frac{\omega}{4}\right) \cdots = \prod_{j=1}^{\infty} H\left(\frac{\omega}{2^j}\right). \quad (2.62)$$

Συνεπώς η λειτουργία κλιμάκωσης είναι απόλυτα καθορισμένη από το βαθυπερατό φίλτρο  $h$ .

Συνεπώς, είναι αναμενόμενο ότι η κατάσταση ορθογωνότητας στο  $(\phi(x-k))_k$  θα έπρεπε να αντικατοπτρίζεται επίσης από το φίλτρο  $h = (h_k)_k$ . Για να επαληθευτεί αυτή η προσδοκία, πρώτα ας καθοριστεί μια περιοδική λειτουργία  $2\pi$

$$A(\omega) = \sum_n |\hat{\phi}|^2(\omega + 2n\pi). \quad (2.63)$$

Από τη σχέση της δυαδικής κλίμακας στον τομέα συχνότητας (2.61), έχουμε

$$\begin{aligned} A(2\omega) &= \sum_n |\hat{\phi}|^2(2\omega + 2n\pi) \\ &= \sum_{n=2m} |H|^2(\omega) |\hat{\phi}|^2(\omega + 2m\pi) + \sum_{n=2m+1} |H|^2(\omega + \pi) |\hat{\phi}|^2(\omega + \pi + 2m\pi). \end{aligned}$$

Επομένως παίρνουμε μια από τις πιο σημαντικές εξισώσεις MRA:

$$A(2\omega) = |H|^2(\omega) A(\omega) + |H|^2(\omega + \pi) A(\omega + \pi). \quad (2.64)$$

Από την άλλη, εφόσον τα  $(\phi(x-k))_k$  είναι ορθογωνικά κάποιος έχει

$$\langle \phi(x), \phi(x-k) \rangle = \delta_k, \quad \text{or equivalently, } \frac{1}{2\pi} \langle \hat{\phi}, e^{-ik\omega} \hat{\phi} \rangle = \delta_k.$$

Ακόμη να δηλωθεί το εσωτερικό προϊόν στο  $L^2(0, 2\pi)$  από  $\langle \cdot, \cdot \rangle$ . Η τελευταία ταυτότητα είναι τότε ισάξια με το

$$\frac{1}{2\pi} \langle A(\omega), e^{-ik\omega} \rangle = \delta_k,$$

το οποίο άμεσα συνεπάγεται ότι

$$A(\omega) \equiv 1 \quad \text{in } L^2(0, 2\pi).$$

Από τη γενική ταυτότητα (2.64) έχουμε εξάγει συνεπώς μια απαραίτητη κατάσταση στο βαθυπερατό φίλτρο  $h = (h_k)_k$  είναι ένα ορθογωνικό MRA:

$$1 = |H|^2(\omega) + |H|^2(\omega + \pi), \quad \omega \in (0, \pi). \quad (2.65)$$

Η βαθυπερατή κατάσταση  $H(0) = 1$  υποδηλώνει άμεσα ότι  $H(\pi) = 0$ , και το ψηφιακό φίλτρο όντως καταστέλλει υψηλές συχνότητες. Στην πράξη η ταυτότητα (2.65) δίνει μια καλή αφετηρία για το σχεδιασμό ορθογωνικού MRA [96].

Μέχρι εδώ έχουμε συζητήσει μόνο για τη λειτουργία κλιμάκωσης  $\varphi$  και το συνδεδεμένο του βαθυπέρατο φίλτρο  $h = (h_k)_k$ . Βεβαίως, η πραγματική ομορφιά του MRA δεν μπορεί να τελειώσει εδώ εφόσον τα κυματίδια δεν έχουν αναδυθεί από την ανάλυση.

Στο MRA, τα κυματίδια αναλογούν στις λεπτομέρειες που χάνονται κατά τη διάρκεια των μεταφορών από σωστές κλίμακες σε πιο άτεχνες. Λόγω της ομοιότητας της δυαδικής κλίμακας στο MRA, αρκεί να εστιάσουμε σε ένα αντιπροσωπευτικό ζευγάρι παρακείμενων κλιμάκων: το  $V_0 \subseteq V_1$ .  $V_1$  έχει καλύτερες λεπτομέρειες και η ορθογωνική προβολή

μιας εικόνας  $u_1 \in V_1$  πάνω στο  $u_0 \in V_0: u_0 = P_0 u_1$  γίνεται για να εξαλείψει ορισμένες λεπτομέρειες του  $u_1$  που είναι αδύνατο να ανιχνευθούν στο  $V_0$ . Έστω ότι το  $W_0$

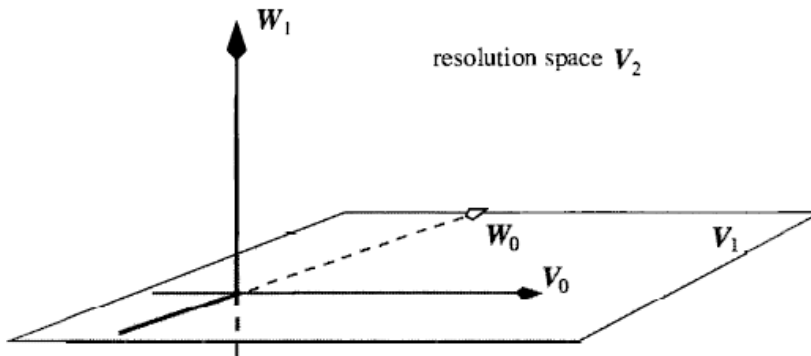
υποδεικνύει την αχτίνα του  $(I - P_0)|_{V_1}$ ,

ή ισοδύναμα, το διάστημα των «λεπτομερειών». Τότε

$$V_1 = V_0 \oplus W_0, \quad P_0 W_0 = \{0\}. \quad (2.66)$$

Γενικά, έστω ότι το  $W_j$  υποδεικνύει το διάστημα των λεπτομερειών στην κλίμακα  $\lambda = 2^{-j}$ , i.e.,  $V_{j+1} = V_j \oplus W_j$ . Τότε είναι εύκολο να δούμε ότι η ομοιότητα της δυαδικής κλίμακας του MRA κληρονομείται πιστά:

$$\eta(x) \in W_j \Leftrightarrow \eta(2x) \in W_{j+1}, \quad j \in \mathbb{Z}.$$



**Εικόνα 2.16.** Το MRA σαν αποσυνθέσεις διαστήματος (Χίλμπερτ):

το καλύτερο διάστημα ευκρίνειας εικόνας  $V_2$  είναι αποσυντεθημένο στο διάστημα της λεπτομέρειας (ή του κυματιδίου)  $W_1$  και το πιο άτεχνο διάστημα ευκρίνειας  $V_1$ . Η ίδια διαδικασία εφαρμόζεται και στο  $V_1$  και σε όλα τα άλλα  $V_j$  [290].

Η πληρότητα τώρα ερμηνεύεται ως



$$L^2(\mathbb{R}) = V_J \oplus \sum_{j \geq J} W_j = \sum_{j=-\infty}^{\infty} W_j, \quad (2.67)$$

όπου το  $J$  είναι οποιοδήποτε αυθαίρετο επίπεδο αναφορικής κλίμακας και τα δυο άπειρα αθροίσματα υποδεικνύουν πάντοτε τα μικρότερα υποδιαστήματα του  $L^2(\mathbb{R})$  τα οποία περιλαμβάνουν όλα τα σχετικά διαστήματα κυματιδίων. Η φόρμουλα δείχνει ότι μια γενική εικόνα  $u(x) \in L^2(\mathbb{R})$  είναι το συνολικό αποτέλεσμα των λεπτομερειών του σε όλες τις κλίμακες. Η Εικόνα 2.16 δίνει μια οπτικοποίηση μικρών διαστάσεων ως προς τις σχέσεις ανάμεσα σε διαφορετικά υποδιαστήματα.

Ας εισάγουμε τον χαρακτηρισμό διαστολής για ένα δεδομένο υποδιάστημα

$$U \subseteq L^2(\mathbb{R}),$$

$$U(2x) = \{f(2x) | f \in U\}.$$

Τότε το (2.67) θα μπορούσε να γραφτεί επίσης ως

$$L^2(\mathbb{R}) = \sum_{j=-\infty}^{\infty} W_0(2^j x),$$

Το οποίο δείχνει ξεκάθαρα ότι το  $W_0$  καθορίζει ολοκληρωτικά την αντιπροσώπευση και ανάλυση των εικόνων στο  $L^2(\mathbb{R})$ .

Για να αποκαλύψουμε περαιτέρω τη δομή του  $W_0$  ας θεωρήσουμε ένα γενικό στοιχείο  $\eta(x) \in W_0$ . Από το (2.66) κάποιος πρέπει να έχει

$$\eta(x) = 2 \sum_k g_k \phi(2x - k) = \sum_k (\sqrt{2} g_k) \phi_{1,k}, \quad g_k = \frac{1}{\sqrt{2}} \langle \eta, \phi_{1,k} \rangle, \quad (2.68)$$

ή στον τομέα συχνότητας

$$\hat{\eta}(\omega) = G\left(\frac{\omega}{2}\right) \hat{\phi}\left(\frac{\omega}{2}\right). \quad (2.69)$$

Η ομάδα των συντελεστών  $g = (g_k)_k$  πρέπει να αντικατοπτρίζουν την φύση των "λεπτομερειών", π.χ., "i.e.,  $W_0 \perp V_0$ , το οποίο ανέρχεται σε  $\langle \eta, \phi(x+k) \rangle = 0, k \in \mathbb{Z}$ , ή στον τομέα συχνότητας

$$\langle (\hat{\eta}\hat{\phi})(\omega), e^{ik\omega} \rangle = 0, \quad k \in \mathbb{Z}.$$

Μετά την περιοδολόγηση του  $2\pi$  οδηγεί στο

$$\sum_n \left( \widehat{\eta\phi} \right) (\omega + 2n\pi) = 0 \quad \text{in } L^2(0, 2\pi).$$

Έχοντας αντικαταστήσει το  $\omega$  με το  $2\omega$  και εφαρμόζοντας τις σχέσεις δυαδικής κλίμακας (2.61) και (2.69), λαμβάνουμε

$$G(\omega)\bar{H}(\omega)A(\omega) + G(\omega + \pi)\bar{H}(\omega + \pi)A(\omega + \pi) = 0,$$

όπου το  $A$  έχει ως δίνεται στο (2.63). Εφόσον  $A(\omega) = 1$  λόγω της ορθογωνιότητας του  $(\varphi(x-k))_k$  έχουμε

$$G(\omega)\bar{H}(\omega) + G(\omega + \pi)\bar{H}(\omega + \pi) = 0. \quad (2.70)$$

$$\lambda(\omega) = G(\omega)e^{iL\omega} / \bar{H}(\omega + \pi)$$

Ορίζουμε το ... για ένα σταθερό περιττό ακέραιο  $L$ . Τότε η τελευταία εξίσωση γίνεται

$$\lambda(\omega) = \lambda(\omega + \pi),$$

Υπονοώντας ότι το  $\lambda(\omega)$  πρέπει να είναι μια περιοδική λειτουργία  $\pi$  ή  $C(\omega) = \lambda(\omega/2)$   $2\pi$ -periodic.

Ορίζουμε ένα πίνακα

$$G_0(\omega) = e^{-i\omega L} \bar{H}(\omega + \pi).$$

Τότε

$$G(\omega) = G_0(\omega)C(2\omega)$$

και

$$\hat{\eta}(\omega) = C(\omega)G_0\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right) = C(\omega)\hat{\psi}(\omega),$$

όπου

$\psi = (G_0(\frac{\omega}{2})\hat{\phi}(\frac{\omega}{2}))^\vee$ , καθορισμένο από τον πίνακα  $G_0$  ονομάζεται το *μητρικό κυματίδιο*. Ας υποθέσουμε ότι οι συντελεστές  $C(\omega)$  είναι  $(c_k)_k$ . Έτσι έχει εδραιωθεί ότι ένα

γενικό στοιχείο λεπτομέρειας  $\eta \in W_0$  πρέπει να είναι υπό τη μορφή

$$\eta(x) = \sum_k c_k \psi(x - k).$$

Συνεπώς, η λεπτομέρεια διαστήματος  $W_0$  καλύπτεται από όλες τις μεταφράσεις των ακέραιων του μητρικού κυματιδίου  $\psi(x)$ .

$$(\psi(x - k))_k$$

Περαιτέρω δείχνουμε ότι το  $(\psi(x - k))_k$  είναι για τη ακρίβεια μια ορθογωνική βάση του  $W_0$ . Όσον αφορά τη λειτουργία κλιμάκωσης ας ορίσουμε

$$B(\omega) = \sum_n |\hat{\psi}|^2(\omega + 2n\pi).$$

$$B(\omega) = 1.$$

Τότε είναι αρκετό να αποδείξουμε ότι  $B(\omega) = 1$ . Παρόμοια με τη λειτουργία κλιμάκωσης η δυαδική σχέση (2.69) οδηγεί εύκολα στο

$$\begin{aligned} B(2\omega) &= |G_0|^2(\omega)A(\omega) + |G_0|^2(\omega + \pi)A(\omega + \pi) \\ &= |H|^2(\omega + \pi)A(\omega) + |H|^2(\omega)A(\omega + \pi) \\ &= |H|^2(\omega + \pi) + |H|^2(\omega) = 1. \end{aligned}$$

Εδώ έχουμε εφαρμόσει το  $A(\omega) \equiv 1$  και το (2.65).

Βάσει ομοιότητας κλίμακας,

$$\{\psi_{j,k} = 2^{j/2}\psi(2^j - k) \mid j, k \in \mathbb{Z}\}$$

είναι μια ορθογωνική βάση για το  $L^2(\mathbb{R})$ , με τη οποία άποψη το  $\psi$  έχει αποκαλεστεί με προτίμηση το μητρικό κυματίδιο. Ικανοποιεί την αποκαλούμενη *εξίσωση κυματιδίου*

$$\psi(x) = 2 \sum_k g_k \phi(2x - k) \quad (2.71)$$

για μια ομάδα κατάλληλων συντελεστών  $g = (g_k)$  όπως συζητήθηκε πιο πάνω. Ας υποθέσουμε ότι το  $G(\omega)$  δηλώνει τον μετασχηματισμό  $g$  του Φουριέ. Τότε η απαραίτητη κατάσταση (2.70), σε συνδυασμό με το  $H(0) = 1$  και το  $H(\pi) = 0$ , οδηγεί στο

$$G(0) = 0 \text{ and } \hat{\psi}(0) = G(0)\hat{\phi}(0) = 0,$$

και συνεπάγεται ότι  $\int_R \psi = 0$ . Αυτές οι προϋποθέσεις στο  $G$  και στο  $\psi$  ονομάζονται οι *υπιπερατές* συνθήκες των κυματιδίων, εφόσον και οι δυο φιλτράρουν στοιχεία χαμηλής συχνότητας κοντά στο  $\omega \approx 0$ . Συνοπτικά, ένα ορθογωνικό MRA κουβαλά τέσσερα κομμάτια πολύ σημαντικών δεδομένων: τη λειτουργία κλιμάκωσης  $\phi$  και το βαθυπερατό της φίλτρο  $h$  συνδεδεμένο με την εξίσωση διαστολής

$$\phi(x) = 2 \sum_k h_k \phi(2x - k) \quad (2.72)$$

και το μητρικό κυματίδιο  $\psi$  και το υπιπερατό φίλτρο του  $g$  συνδεδεμένο με τη λειτουργία κλιμάκωσης μέσω της εξίσωσης κυματιδίου

$$\psi(x) = 2 \sum_k g_k \phi(2x - k). \quad (2.73)$$

Για την ορθογωνιότητα, τα δυο φίλτρα πρέπει να ικανοποιούν

$$1 = |H|^2(\omega) + |H|^2(\omega + \pi), \quad (2.74)$$

$$0 = (G\bar{H})(\omega) + (G\bar{H})(\omega + \pi). \quad (2.75)$$

Το ακριβές σχέδιο του ορθογωνικού MRA αρχίζει από αυτές τις δυο τελευταίες εξισώσεις στο βαθυπερατό φίλτρο  $h$  και στο υπιπερατό φίλτρο  $g$ . Στις περισσότερες εφαρμογές επεξεργασίας σήματος και εικόνας τόσο το  $h$  όσο και το  $g$  περιέχουν ένα πεπερασμένο αριθμό μη-μηδενικών συντελεστών· τότε ουσιαστικά γίνεται ένα σχεδιαστικό πρόβλημα πολυώνυμου (ή πολυώνυμου του Λωρεντ) [95, 96, 290]. Τέτοια φίλτρα πιστεύεται ότι έχουν αποκρίσεις πεπερασμένου παλμού (ΑΠΠ) στην επεξεργασία ψηφιακού σήματος [237]. Η εικόνα 2.17 δείχνει ένα ζευγάρι λειτουργίας κλιμάκωσης και μητρικού κυματιδίου.

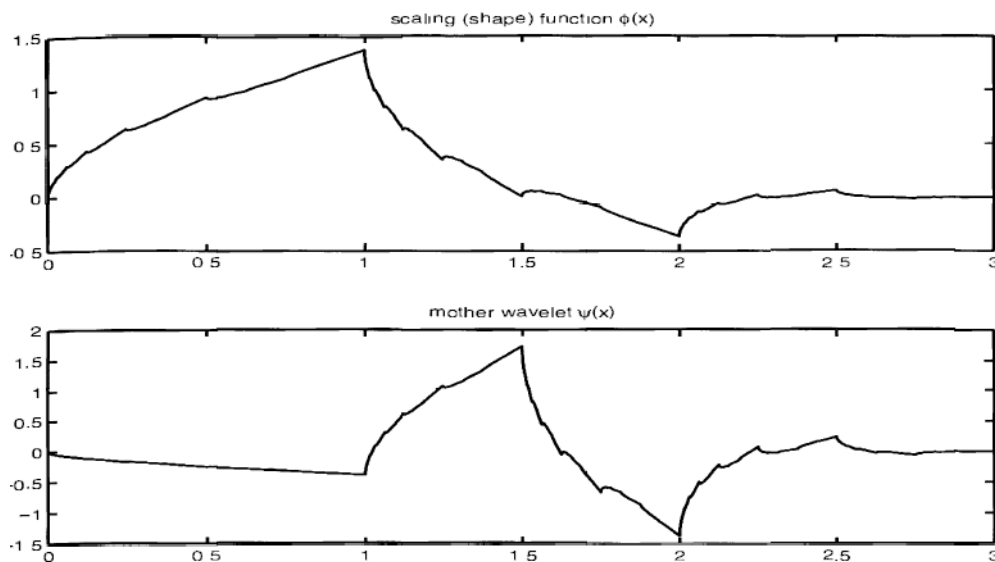
## 2.6.6 Γρήγορη Ανάλυση και Σύνθεση Εικόνων μέσω Συστοιχιών Φίλτρων

Τώρα εξηγούμε περαιτέρω πως το πλαίσιο του MRA οδηγεί σε αποτελεσματικούς υπολογιστικούς αλγορίθμους για ανάλυση και σύνθεση εικόνων πολλαπλής σύνθεσης.

Από την τοποθέτηση πληρότητας, οποιαδήποτε εικόνα  $u(x) \in L^2(\mathbb{R})$  θα μπορούσε να υπολογιστεί σε οποιαδήποτε ακρίβεια από τις προβολές της  $u_j = P_j u$  πάνω σε αυξημένες ευκρίνειες  $V_j$ :

$$u_j = \sum_k c_k^j \phi_{j,k}, \quad c_k^j = \langle u, \phi_{j,k} \rangle. \quad (2.76)$$

Η αποτελεσματικότητα του MRA είναι σε μεγάλο βαθμό αποτέλεσμα των ενδογενών συνδέσεων τέτοιων επεκτάσεων σε διαφορετικές κλίμακες, όπως καλλιεργούνται από την τοποθέτηση δυαδικής ομοιότητας στο σχεδιασμό του MRA. Έτσι σε περιπτώσεις όπου πολλαπλοί χρήστες ενδιαφέρονται για διαφορετικές διεργασίες και κλίμακες, υπάρχει μια γενική δομή αποθήκευσης και ανάκτησης δεδομένων η οποία θα μπορούσε να ικανοποιήσει τις πλείστες ανάγκες αποτελεσματικά.



**Εικόνα 2.17.** Ένα ζευγάρι συμπαγώς στηριζόμενης λειτουργίας κλιμάκωσης  $\phi(x)$  και μητρικού κυματιδίου  $\psi(x)$  από το σχέδιο του Ντομπεσι [96].

Συγκεκριμένα, γίνεται αχρείαστο να επεξεργαστούμε τα ακριβά ολοκληρώματα για όλα τα  $j$  και  $k$  στο (2.76).

Τέτοια αποτελεσματικότητα πηγάζει από τις δυο κύριες εξισώσεις (2.72) και (2.73). Από την τοποθέτηση δυαδικής ομοιότητας, κάποιος έχει

$$\phi_{j,k} = \sum_l h_l \phi_{j+1,2k+l} \text{ and } \psi_{j,k} = \sum_l g_l \phi_{j+1,2k+l}.$$

$$u_{j+1} = u_j + w_j \in V_j \oplus W_j.$$

Ας υποθέσουμε ότι . Τότε

$$c_k^j = \langle u, \phi_{j,k} \rangle = \left\langle u, \sum_l h_l \phi_{j+1,2k+l} \right\rangle = \sum_l h_l c_{2k+l}^{j+1}.$$

Παρομοίως και για τους συντελεστές κυματιδίων,

$$d_k^j = \langle u, \psi_{j,k} \rangle = \sum_l g_l c_{2k+l}^{j+1}.$$

Ας ορίσουμε τις μεταφορές των δυο φίλτρα (με πραγματική αξία) από χωρική αντιστροφή:

$$\tilde{g}_l = g_{-l} \text{ and } \tilde{h}_l = h_{-l},$$

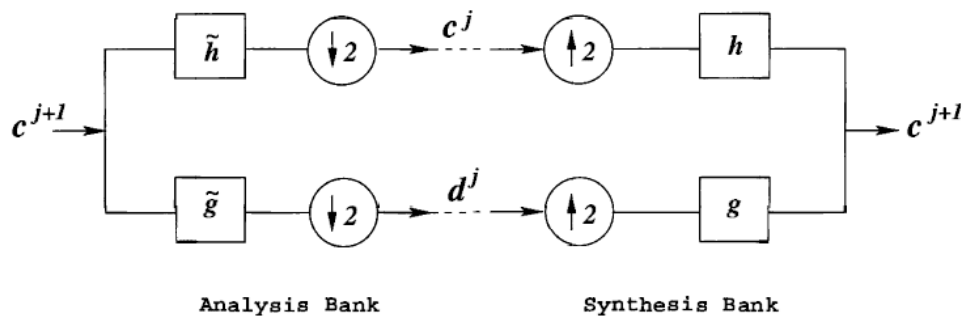
και όπως και στην επεξεργασία ψηφιακού σήματος ορίζεται επίσης ο χειριστής υποδειγματοληψίας στο  $\hat{f}$ :

$$(\downarrow 2)c = (\downarrow 2)(\dots, c_{-2}, c_{-1}, c_0, c_1, c_2, \dots) = (\dots, c_{-2}, c_0, c_2, \dots) \quad (2.77)$$

α φέρνοντας όλα τα περιττά στοιχεία. Τότε τα πιο πάνω αποτελέσματα μπορούν να εκφραστούν συνοπτικά από το

$$c^j = (\downarrow 2)(\tilde{h} * c^{j+1}) \text{ and } d^j = (\downarrow 2)(\tilde{g} * c^{j+1}), \quad (2.78)$$

όπου τα σύμβολα αστερίσκων υποδηλώνουν διακριτικές συγκεράσεις ακολουθιών.



**Εικόνα 2.18.** Γρήγορος μετασχηματισμός κυματιδίων μέσω συστοιχιών φίλτρων: η συστοιχία ανάλυσης διπλού καναλιού (ή αποσύνθεσης) και σύνθεσης (ή επαναδόμησης).

Είναι ακριβώς εδώ που το μαθηματικό πλαίσιο του MRA συναντά απλά την σύγχρονη επεξεργασία ψηφιακών σημάτων (ΕΨΣ). Στην ΕΨΣ, αυτές οι δυο ταυτότητες οδηγούν σε μια

ανάλυση συστοιχίας φίλτρου  $(\tilde{h}, \tilde{g})$ , διπλού καναλιού, η οποία παίρνει το  $c^{j+1}$  σαν είσοδο και το  $(c^j, d^j)$  σαν έξοδο.

Η επανάληψη αυτής του απλού κεφαλαίου ΕΨΣ οδηγεί σε γρήγορη αποσύνθεση κυματιδίων: ξεκινώντας από οποιαδήποτε καλή κλίμακα  $\lambda = 2^{-j}$ ,

$$c^j \rightarrow (d^{j-1}, c^{j-1}) \rightarrow \dots \rightarrow (d^{j-1}, d^{j-2}, \dots, d^0, c^0). \quad (2.79)$$

Επομένως, όταν η αντιπροσώπευση σε μια καλή κλίμακα είναι διαθέσιμη, όλες οι αποσυνθέσεις των πιο άτεχνων κλιμάκων μπορούν να εξαχθούν απλά εφαρμόζοντας την ανάλυση συστοιχίας διπλού καναλιού, η οποία μας χαρίζει τον ακριβό υπολογισμό όλων των ολοκληρωμάτων των εσωτερικών προϊόντων.

Η συστοιχία ανάλυσης (στην Εικόνα 2.18) συνοδεύεται από τη συνοδευτική σύνθεση ή τη συστοιχία επαναδόμησης. Στο ορθογωνικό πλαίσιο MRA, η συστοιχία σύνθεσης είναι η μετατόπιση της συστοιχίας ανάλυσης, πολύ όμοια στο πασίγνωστο γεγονός της γραμμικής άλγεβρας σύμφωνα με το οποίο αντιστρέφοντας ένας ορθογώνιος πίνακας αναλογεί στη

μετατόπιση του  $(\tilde{h}, \tilde{g})$  ζευγαριού ανάλυσης της συστοιχίας είναι απλά  $(h, g)$ , το αρχικό ζευγάρι του φίλτρου στη διαστολή και τις εξισώσεις του κυματιδίου.

Η μετατόπιση του χειριστή της υποδειγματοληψίας  $(\downarrow 2)$ , όπως είναι επίσης γνωστό στις μεθόδους πολλαπλών πλεγμάτων [34], είναι ο χειριστής αύξησης της δειγματοληψίας  $(\uparrow 2)$ :

$$(\uparrow 2)c = (\uparrow 2)(\dots, c_{-1}, c_0, c_1, \dots) = (\dots, c_{-1}, 0, c_0, 0, c_1, 0, \dots).$$

Τότε η φόρμουλα σύνθεσης δίνεται από το (δες Εικόνας 2.18)

$$c^{j+1} = h * (\uparrow 2)(c^j) + g * (\uparrow 2)(d^j).$$

Σε συνδυασμό, το MRA γίνεται εντελώς ψηφιακό και στο ενδιάμεσο η ανάλυση συστοιχίας και η σύνθεση συστοιχίας στην Εικόνα 2.18, π.χ. στην περιοχή κυματιδίων επιτρέπει τη μεγάλη ελευθερία για κωδικοποίηση εικόνων, μετάδοση, επεξεργασία, ή ανάλυση.

Η δομή συστοιχίας φίλτρου εισηγείται επίσης αμέσως τον τρόπο σχεδιασμού του

αποκαλούμενου διορθογωνικού MRA στο οποίο το  $(\psi_{j,k})_{j,k}$  είναι μόνο μια βάση του Ριεζ για  $L^2(\mathbb{R})$  και το  $(\phi_{j,k})_{j,k}$  για το  $V_j$ . Ψηφιακά, σημαίνει ότι το ζευγάρι σύνθεσης φίλτρου

$(h_s, g_s)$  δεν απαιτείται να είναι πλέον η μετατόπιση του ζευγαριού ανάλυσης

$(h_a, g_a)$ . Αντί γι'αυτού, ένα διορθογωνικό MRA επιβάλλει μόνο την προϋπόθεση της τέλειας ή χωρίς απώλεια επαναδόμησης:

$$(h_s^*)(\uparrow 2)(\downarrow 2)(h_a^*) + (g_s^*)(\uparrow 2)(\downarrow 2)(g_a^*) = \text{Id}, \quad \text{the identity operator.}$$

Φυσικά, όπως υπολογίζοντας την αντιστροφή ενός γενικού συγγενικού πίνακα  $A$  στην γραμμική άλγεβρα, κάποιος πρέπει να βεβαιωθεί ότι οι συστοιχίες φίλτρων είναι σε καλή κατάσταση, έτσι ώστε τόσο τα βήματα ανάλυσης όσο και τα βήματα σύνθεσης είναι σταθερά. Η ιδέα της διορθογωνικότητας διευρύνει κατά πολύ την οικογένεια κυματιδίων που βασίζονται στο MRA. Και ακόμη πιο αξιοσημείωτο, επιτρέπει στις λειτουργίες κλιμάκωσης, τα κυματίδια και τα φίλτρα τους να είναι συμμετρικά ή να έχουν ίσες φάσεις [290], μια καίρια τεχνική αξία στα σήματα και κυρίως στην επεξεργασία εικόνων.

Βέβαια, το  $(p_\nu)_\nu$  σαν κατανομή πιθανότητας πρέπει επίσης να ικανοποιεί τις φυσικές συνθήκες μη-αρνητικότητας και ενότητας. Εισάγοντας τους πολλαπλασιαστές Λαγκρανς  $\beta$  για το  $E$ ,  $\mu$  για  $n$ , και  $\delta$  για την κανονικοποίηση ενότητας, οδηγούμαστε στο πρόβλημα αβίαστης βελτιστοποίησης

$$\max_{(p_\nu)} \left( - \sum_\nu p_\nu \ln p_\nu + \beta \sum_\nu E_\nu p_\nu + \mu \sum_\nu n_\nu p_\nu + \delta \sum_\nu p_\nu \right). \quad \text{Πέρνον}$$

τας τα παράγωγα κατά μήκος κάθε  $p_\nu$  δίνει

$$- \ln p_\nu - 1 + \beta E_\nu + \mu n_\nu + \delta = 0, \quad \nu = 0, 1, \dots,$$

οποίο δηλώνει ότι

$$p_\nu = \frac{1}{Z_{\beta, \mu}} e^{-\beta E_\nu - \mu n_\nu}, \quad \nu = 0, 1, \dots,$$



όπου η λειτουργία διαίρεσης  $Z$  (αρχικά συνεισφερόμενη από το  $\delta$ ) έχει συνδεθεί στο  $\beta$  και  $\mu$  για την κανονικοποίηση πιθανότητας:

$$Z_{\beta, \mu} = \sum_{\nu} e^{-\beta E_{\nu} - \mu n_{\nu}}.$$

Αυτός είναι ο νόμος κατανομής του Γκιμπς για τα GCE. Χωρίς ανταλλαγές σωματιδίων, αναπαράγει το νόμο του Γκιμπς (2.29) για τα CE.

Αυτή η κομψή προσέγγιση στη στατιστική μηχανική έχει εμπνεύσει κατά πολύ το γενικό πλαίσιο του μοτίβου ανάλυσης εικόνας που αναπτύχθηκε από τους Γκέμαν και Γκέμαν, Μάμφορτ και Ζου και άλλους. Το νέο και προκλητικό υλικό για την ανάλυση εικόνας είναι και πάλι να ορίσουμε κατάλληλα οπτικά βαρυσήμαντα χαρακτηριστικά, εκείνα τα οποία είναι παράλληλα τόσο στις εκτατικές όσο και στις εντατικές μεταβλητές στη θερμοδυναμική και στη στατιστική μηχανική.

## 2.4 Η Μπεϋζιανή Στατιστική Επαγωγή

Αυτή η ενότητα επεξηγεί σε συντομία το βασικό ρόλο της θεωρίας της Μπεϋζιανής λογικής στην ανάλυση εικόνας και όρασης.

### 2.4.1 Επεξεργασία Εικόνας ή Οπτική Αντίληψη ως Επαγωγή

Η επεξεργασία εικόνας ή η οπτική αντίληψη μπορεί να διατυπωθεί συνοπτικά σαν ένα πρόβλημα επαγωγής: έχοντας τα μελετημένα δεδομένα της εικόνας  $Q$ , καθορίζει τα μοτίβα κλειδιά ή χαρακτηριστικά  $F$  ενσωματωμένο εντός. Αν τα δεδομένα εικόνας  $Q$  θεωρούνται, τουλάχιστο εννοιολογικά, ως τα δείγματα στοχαστικών διανομών που είναι είτε παραμετροποιημένες από ή δημιουργημένες από το μεταβλητό χαρακτηριστικό  $F$ , τέτοια απόφαση ανήκει φυσικά στο βασίλειο της στατιστικής επαγωγής. Εδώ τόσο το  $Q$  όσο και το  $F$  μπορούν να περιέχουν πολλαπλούς συνδέσμους π.χ. το  $Q$  υποδηλώνει μια ακολουθία εικόνων σε ένα τμήμα βίντεο ή από τη συνεχόμενη σπινθηρογραφία ενός ασθενούς.

Πρώτα, ας υποθέσουμε ότι το  $F$  είναι ένα ανοιχτό-ή-κλειστό δυαδικό χαρακτηριστικό, δηλαδή,

συμβολίζει ένα συγκεκριμένο χαρακτηριστικό που είναι παρόν ή απόν. Έστω ότι  $F = 0$  και 1 υποδηλώνουν ότι το χαρακτηριστικό είναι ανοιχτό και κλειστό ξεχωριστά. Τότε η επαγωγή στις αξίες του  $F$  δεδομένου του  $Q$  μπορούν να επιτευχθούν από τον έλεγχο σημαντικότητας.

Ένας έλεγχος σημαντικότητας οδηγεί σε μια επιλογή ανάμεσα στην αποκαλούμενη μηδενική υπόθεση (δηλαδή  $F = 0$ ), υποθέτοντας ότι το στοχευόμενο χαρακτηριστικό είναι παρόν, και τις εναλλακτικές

(δηλαδή  $F = 1$ ), βασισμένο σε κάποια στατιστική κριτηρίων  $T = T(Q | F = 0)$  και

το συνδεδεμένο του κατώφλιο  $t_c$ . Δεδομένου ενός συγκεκριμένου δείγματος  $Q = q$  των δεδομένων εικόνας, ο κανόνας απόφασης ενός ελέγχου σημαντικότητας θα μπορούσε να είναι

$$T(q | F = 0) < t_c \Rightarrow \text{accept } F = 0; \quad \text{otherwise, } \Rightarrow \text{accept } F = 1.$$

Συνεπώ

ς το  $t_c$  αποκαλείται συχνά το όριο απόφασης.

Για παράδειγμα, ας υποθέσουμε ότι το  $Q$  υποδηλώνει την ελαφριά πυκνότητα σε κάποιο σταθερό μικροστοιχείο εικόνας και το  $F = 0$  αναλογεί στην υπόθεση ότι το  $Q$  δημιουργείται από ένα Γκαουσιανό

$N(\mu = 90, \sigma^2 = 100)$ , ) ενώ το  $F = 1$  σε διαφορετική περίπτωση. Τότε το ζευγάρι

$$(\mu = 90, \sigma = 10)$$

είναι το μοτίβο στόχος στο παρόν σενάριο. Ας ορίσουμε τη στατιστική κριτηρίων να είναι

$$T_0 = \frac{Q - \mu}{\sigma} = \frac{Q - 90}{10} \quad \text{and } T = |T_0|.$$

Η σημαντική απόφαση ορίου  $t_c$  θα μπορούσε να είναι για παράδειγμα, το 95<sup>ο</sup> συμμετρικό ποσοστό, δηλαδή η μοναδική αξία  $t_c$  έτσι ώστε

$$\text{Prob}(|T_0| \geq t_c) = 5\% = 0.05, \quad \text{if indeed } T_0 \sim N(0, 1).$$

Με αυτό το κριτήριο, η πιθανότητα *αστόχησης* (δηλαδή αποτυχίας να ανιχνεύσουμε το μοτίβο στόχο ενώ είναι για την ακρίβεια παρόν) ελέγχεται κάτω από το 5 τις εκατό.

Για ένα πιο γενικό διαγωνισμό ανάμεσα σε πολλαπλές αποκλειστικές υποθέσεις

$$F = f_1, f_2, \dots, \text{ or } f_N,$$

μια άλλη δημοφιλής μέθοδος επαγωγής είναι το κριτήριο της *μέγιστης πιθανότητας* (ΜΠ). Υπό την αρχή της ΜΠ, δεδομένων οποιωνδήποτε μελετημένων δεδομένων εικόνας  $Q = q$ , το κρυμμένο χαρακτηριστικό υπολογίζεται από

$$\hat{F} = \underset{f_i \in f_{1N}}{\text{argmax}} p(Q = q | F = f_i);$$

δηλαδή το ενσωματωμένο χαρακτηριστικό αναγνωρίζεται με αυτό που οδηγεί στη ΜΠ των μελετημένων δεδομένων εικόνας.

Για παράδειγμα, ας υποθέσουμε ότι μια εικόνα χαμηλής ποιότητας  $Q = u_0$  υπό μια διόπτρα νυχτερινής παρατήρησης φωτογραφίζει ένα πεζό άντρα  $F$  με καπέλο του μπέϊσμπολ και ότι αυτός ο

άντρας είναι γνωστός ως ένας από τους δίδυμους  $f_1 = \text{Mike and } f_2 = \text{Jack}$ . Ο Μάικ και ο Τζακ έχουν σχεδόν την ίδια σωματική εμφάνιση, καθώς και ρούχα, μόνο που ο Μάικ αγαπά τα

καπέλα του μπέϊσμπολ ενώ ο Τζακ τα φορά μόνο περιστασιακά. Σαν ένα μοντέλο, για παράδειγμα, ας υποθέσουμε ότι

$$p(u_0 | F = f_1) = p(\text{wearing a hat} | \text{Mike}) = 50\%,$$

$$p(u_0 | F = f_2) = p(\text{wearing a hat} | \text{Jack}) = 10\%,$$

εννοώντας από την άποψη της οριακής συχνότητας ο Μάικ φορά ένα καπέλο του μπέϊσμπολ μέρα παρά μέρα ενώ ο Τζακ το φορά μόνο τρεις φορές το μήν. Έτσι σύμφωνα με το κριτήριο ΜΠ, το άτομο

(ή μοτίβο) στην εικόνα είναι  $\hat{F} = f_1 = \text{Mike}$ .

#### 2.4.4 Η Μπεϋζιανή Επαγωγή: Πόλωση Λόγω Προηγούμενης Γνώσης

Τέτοια απόφαση μπορεί να γίνει πιο εμπλεκόμενη αν άλλες πληροφορίες είναι διαθέσιμες σχετικά με τα μοτίβα στόχου.

Ακολουθώντας το προηγούμενο παράδειγμα ας υποθέσουμε επιπλέον ότι οι δίδυμοι ζουν στη γεννέτηρα τους τη Βοστώνη και ότι

(α) η συγκεκριμένη εικόνα  $Q = u_0$  πάρθηκε στο Παρίσι, στη Γαλλία και

(β) ο Μάικ ως δάσκαλος σε σχολείο ταξιδεύει μόνο περιστασιακά, ενώ ο Τζακ σαν πωλητής σε μια πασίγνωστη διεθνή εταιρεία ταξιδεύει πολύ πιο συχνά.

Για να γίνουν τα πράγματα πιο εύκολα, ας υποθέσουμε ότι οι δίδυμοι ταξιδεύουν μόνο μεταξύ Βοστώνης και Παρισιού και ότι

$$p(F = f_1 | \text{in Paris}) = p(F = \text{Mike} | \text{one of the twins in Paris}) = 10\%,$$

$$p(F = f_2 | \text{in Paris}) = p(F = \text{Jack} | \text{one of the twins in Paris}) = 90\%.$$

Εφόσον η φωτογραφία  $Q = u_0$  είναι γνωστό ότι πάρθηκε στο Παρίσι, κάποιος θα μπορούσε απλά να αφήσει το υποθετικό σύμβολο και γράφει

$$p(F = f_1) = 10\% \text{ and } p(F = f_2) = 90\%.$$

Με αυτή τη νέα πληροφορία, η μέθοδος ΜΠ που εισήχθηκε στην προηγούμενη ενότητα γίνεται προφανώς ανεπαρκής. Ας θεωρήσουμε τη φανταστική ακραία περίπτωση όταν

$$p(F = f_1) = 0 \text{ and } p(F = f_2) = 100\%,$$

εννοώντας ότι ο Μάικ δεν φεύγει ποτέ από τη Βοστώνη. Τότε γίνεται παράλογο να αναγνωρίσουμε τον άνθρωπο στην εικόνα  $u_0$  ως τον Μάικ, όπως μόλις συμπεράθηκε προηγουμένως από την αρχή του ΜΠ. Μια πιο λογική στρατηγική επαγωγής πρέπει να επωφεληθεί από τη νέα πληροφορία  $p(F = f)$ .

Η διανομή  $p(F = f)$  λέγεται να είναι η *προ* πληροφορία εφόσον είναι διαθέσιμη *πριν* από την εξασφάλιση του δεδομένου  $Q = u_0$  - η εικόνα  $u_0$  πάρθηκε στο Παρίσι στην περίπτωση μας. Με άλλα λόγια, οι προηγούμενες πιθανότητες του Μάικ και του Τζακ στο Παρίσι δεν επηρεάζονται από το κατά πόσο μια φωτογραφία πάρθηκε στο Παρίσι ή όχι. (Οι *υποθετικές* ή *οπίσθιες* πιθανότητες είναι παρόλα αυτά όντως επηρεασμένες.)

Συνεπώς, από άποψης στοχαστικής μοντελοποίησης, η διανομή του  $p(F = f)$  του μοτίβου στόχου ή χαρακτηριστικού  $F$  αποκαλείται το *αρχικό μοντέλο*, καθώς η πιθανότητα διανομής  $p(Q = u_0 | F = f)$  κεντρική για τη διαδικασία επαγωγής ΜΠ αποκαλείται το *μοντέλο δεδομένων*, ή *μοντέλο παραγωγής δεδομένων* για να δοθεί έμφαση στο γεγονός ότι το  $F$  παράγει

$$Q: F \rightarrow Q.$$

Με τη διαθέσιμη προηγούμενη γνώση, ένα λογικό σχήμα επαγωγής θα έπρεπε να βασίζεται τόσο στο προηγούμενο μοντέλο όσο και στο μοντέλο δεδομένων. Το κεντρικό θέμα είναι πώς να συνδυαστούν αυτά τα δυο κατάλληλα και μια βολική απάντηση θα ήταν να ζυγίσουμε απλά το προϊόν τους:

$$g(F = f | Q = q) = \text{prior model} \times \text{data model} = p(F = f) \times p(Q = q | F = f). \quad (2.33) \quad \text{Κάποιο}$$

ς τότε αναγνωρίζει το μοτίβο στόχο σαν εκείνο που μεγιστοποιεί αυτό το προϊόν:

$$\hat{F} = \underset{f \in f_1 \cup f_2}{\operatorname{argmax}} g(F = f | Q = q).$$

Σε σύγκριση με την προηγούμενη ενότητα, αυτή η νέα στρατηγική επαγωγής μπορεί να ονομαστεί η μέθοδος *μεροληπτικής* ΜΠ και κατά ακρίβεια είναι ακριβώς η δοξασμένη προσέγγιση της Μπεϋζιανής επαγωγής. Συνεχίζοντας στο πιο πάνω πείραμα στον Μάικ και τον Τζακ, η Μπεϋζιανή επαγωγή συγκρίνει τη δύναμη του

$$g(F = f_1 | Q = u_0) = 50\% \times 10\% = 5\% \text{ and } g(F = f_2 | Q = u_0) = 10\% \times 90\% = 9\% \text{ και}$$

φτάνει στο συμπέρασμα ότι ο άνθρωπος με το καπέλο του μπέιζμπολ στην εικόνα  $u_0$  είναι (πιο πιθανόν) ο Τζακ, το οποίο είναι εντελώς αντίθετο στον αρχικό υπολογικό ΜΠ.

Η εγκυρότητα του τύπου  $g(F = f | Q = q)$  του προϊόντος στο (2.33) σαν ένας κατάλληλος τρόπος να συνδυαστούν τόσο το αρχικό μοντέλο όσο και το μοντέλο δεδομένων υποστηρίζεται από τη Μπεύζιανή φόρμουλα για υποθετικές πιθανότητες:

$$p(F = f | Q = q) = \frac{p(F = f)p(Q = q | F = f)}{p(Q = q)},$$

ή ακόμα πιο γενικά, για οποιαδήποτε δυο τυχαία γεγονότα  $A$  και  $B$ ,

$$p(A | B) = \frac{p(A)p(B | A)}{p(B)}.$$

Συνεπώς, ο τύπος  $g$  του προϊόντος είναι ακριβώς η οπίσθια πιθανότητα

$p(F = f | Q = q)$  μέχρι τον παράγοντα παρονομαστή  $p(Q = q)$ , ο οποίος είναι παρόλα αυτά, μια σταθερά όταν επιτευχθεί μια συγκεκριμένη παρατήρηση  $Q$ . Όντως ο παρονομαστής είναι απλώς μια σταθερά κανονικοποίησης πιθανότητας:

$$p(Q = q) = \sum_{f \in \mathcal{F}_1} p(F = f)p(Q = q | F = f).$$

Σε σύγκριση με τη μέθοδο ΜΠ στην προηγούμενη ενότητα, η Μπεύζιανή προσέγγιση μεγιστοποιεί συνεπώς την οπίσθια πιθανότητα, ή εξίσου, ανιχνεύει το πιο πιθανό μοτίβο στόχο  $F = f$  που βασίζεται (ή εξαρτάται) στα εξασφαλισμένα δεδομένα εικόνας  $Q = q$ .

Η θεωρητική βάση – πληροφορία [93] για την Μπεύζιανή προσέγγιση μπορεί να γίνει κατανοητή ως

ακολουθώς. Τα μελετημένο δεδομένο  $Q$  πιστεύεται ότι περιέχει κάποιες πληροφορίες για τον στόχο  $F$ , και συνεπώς το  $F$  θα έπρεπε να είναι αναγνώσιμο ή ανιχνεύσιμο από το  $Q$  με κάποιο άριστο τρόπο. Στην περίπτωση όταν το  $Q$  δεν περιέχει πληροφορίες για το  $F$ , τα δεδομένα γίνονται άχρηστα, και η Μπεύζιανή επαγωγή στο  $F$  εκφυλίζεται σε μια απόλυτη μεγιστοποίηση της αρχικής διανομής

$F$ . Οι πιο πάνω ενστικτώδεις έννοιες της «πληροφορίας» ή της «αμοιβαίας πληροφορίας» μπορεί να ποσοτικοποιηθεί σωστά από τη θεωρία της πληροφορίας [93]. Για παράδειγμα, η αμοιβαία

πληροφορία  $I(F, Q)$  του  $F$  και του  $Q$  καθορίζεται ως

$$I(F, Q) = H(F) - H(F | Q) = H(Q) - H(Q | F) = H(F) + H(Q) - H(F, Q), \quad \text{όπου}$$

$H(F)$ ,  $H(Q)$ , και  $H(F, Q)$  είναι ξεχωριστές ή ενωμένες εντροπίες που καθορίζονται όπως και στην προηγούμενη ενότητα στη θερμοδυναμική και στατιστική μηχανική καθώς το

$H(F | Q)$  και το  $H(Q | F)$  είναι υποθετικές εντροπίες. Αναφέρουμε τον αναγνώστη στο [93] για περαιτέρω πληροφορίες σε αυτές τις έννοιες.

### 2.4.3 Μπεϋζιανή Μέθοδος στην Επεξεργασία Εικόνας

Η Μπεϋζιανή μέθοδος έχει παίξει κεντρικό ρόλο διαμέσου ολόκληρου του φάσματος επεξεργασίας εικόνας και οπτικής αντίληψης. Δείτε, για παράδειγμα, τη μονογραφία *Αντίληψη σαν Μπεϋζιανή Επαγωγή* που συντάχθηκε από τους Νιλ και Ρίτσαρτς [176].

Αναφορικά με τα δύο κτιριακά μπλοκ ενός γενικού μοντέλου Μπεϋζιανής επαγωγής, το αρχικό

μοντέλο  $p(F = f)$  και το μοντέλο δεδομένων  $p(Q = a | F = f)$ , το πρώτο συνήθως επιβάλλει μεγάλες προκλήσεις και συνπώς μένει στην καρδιά πολλών εργασιών μοντελοποίησης στη ανάλυση εικόνας και όρασης.

Από την άποψη επεξεργασίας εικόνας, η δόμηση αρχικών μοντέλων ήχου εικόνας στην ουσία μετρά ως ο σωστός ορισμός της ακριβής σημασίας των «εικόνων», το οποίο ο δεύτερος συγγραφέας ορίζει ως το *κύριο πρόβλημα* της επεξεργασίας εικόνας. Οι πλείστες εργασίες υπολογισμού εικόνας είναι συνήθως κακώς – τειθέμενες, και είναι συχνά η αρχική γνώση στις εικόνες που εισάγει την απαραίτητη απόκλιση, καθώς επίσης μειώνει σημαντικά το μέγεθος του συνόλου των ευποψήφινων εικόνων.

Υπάρχουν γενικά δύο τρόποι δόμησης αρχικών μοντέλων εικόνας: μαθαίνοντας ή μοντελοποιώντας άμεσα.

Με μεγάλες τράπεζες δεδομένων εικόνας που είναι εύκολα διαθέσιμα σε αυτή την περιοχή πληροφοριών, είναι πιθανό να μάθουμε τα κοινά χαρακτηριστικά που είναι ενσωματωμένα σε γενικές εικόνες βασισμένες σε στατιστικές γνώσεις. Για παράδειγμα, οι Ζου και Μάμφορτ [328], και η Ζου, Γου και Μάμφορτ [329] συνδύασαν την αρχή της μέγιστης εντροπίας με κατάλληλα οπτικά φίλτρα για να μάθουμε τις κυρίαρχες στατιστικές φυσικών εικόνων. Παρόλο το μεγάλο αναγνωρισμένο κεντρικό ρόλο της σε ένα αριθμό πειθαρχιών, η θεωρία μάθησης παραμένει ένας νέος τομέας και χρειάζεται πολύ περισσότερη προσοχή και προσπάθειες από σύγχρονους επιστήμονες. Παραπέμπουμε τον αναγνώστη σε ένα πρόσφατο ερευνητικό άρθρο των Πότζιο και Σμείλ [252] καθώς και την κλασσική μονογραφία του Βάπνικ [307] για μια καλή σύνοψη στη θεωρία μάθησης.

Μια λιγότερο υπολογιστική ακριβή προσέγγιση στα αρχικά μοντέλα εικόνας είναι από την άμεση δόμηση, η οποία τυπικά βασίζεται είτε σε κάποιες κοινές οπτικές εμπειρίες ή κάποιες σχετικές αρχές φυσικής για τυποποίηση εικόνας.

Για παράδειγμα, για ένα σύνολο εικόνων που δημιουργήθηκαν τυχαία από δυφασικές συσκευές κρυσταλλικής απεικόνισης, το μοντέλο πλέγματος του Ιζικ στη στατιστική μηχανική προσφέρει ένα αρχικό μοντέλο φυσικής εικόνας [82]. Πιο γενικά αρχικά μοντέλα εικόνας μπορούν να δομηθούν όμοια χρησιμοποιώντας, για παράδειγμα, τα σύνολα του Γκιπς και τους τυχαίους τομείς του Μάρκοβ, όπως μελετήθηκε στη πασίγνωστη δουλειά των Γκέμαν και Γκέμαν [130]. Από την άλλη, εμπνευσμένοι από κάποιες κοινές εμπειρίες στην εικονική αντίληψη, συμπεριλαμβανομένου του μεταφραστικού αναλλοίωτου και του αναλλοίωτου της κλίμακας, οι Μάμφορτ και Γκίτας πρότειναν αρχικά μοντέλα εικόνας βασισμένα σε ορισμένα θεμελιώδη αξιώματα [225].

Λιγότερο απαιτητικά υπολογιστικά σε σύγκριση με την προσέγγιση μάθησης, τα αρχικά μοντέλα εικόνας από μια τέτοια άμεση δόμηση ξεκάθαρα στερούνται προσαρμοστικότητας και πολύ συχνά μπορούν να εφαρμοστούν μόνο ως υπολογισμοί χαμηλής τάξης.

Το επόμενο κεφάλαιο θα αποκαλύψει περισσότερες πληροφορίες στις προσεγγίσεις δόμησης αρχικών μοντέλων εικόνας. Τα υπόλοιπα κεφάλαια αποδεικνύουν περαιτέρω τους σημαντικούς ρόλους τέτοιων αρχικών πληροφοριών εικόνας σε πολλές περιοχές επεξεργασίας εικόνας.

## 2.5 Γραμμικό και Μη Γραμμικό Φιλτράρισμα και Διάχυση

Σε αυτή την ενότητα, εξερευνούμε τις κύριες ιδέες που υποκείμενες σε ένα από τα πιο δυνατά εργαλεία στην ανάλυση - φιλτράρισμα εικόνας και όρασης. Μαθηματικά, το φιλτράρισμα είναι στενά συνδεδεμένο στον τελεστή μάλαξης στην πραγματική ανάλυση και το φαινόμενο διάχυσης στη θεωρία της Μερικής Διαφορικής Εξίσωσης (ΜΔΕ). Στοχαστικά, το φιλτράρισμα μπορεί να επιτευχθεί από τυχαία βαδίσματα μικροσκοπικών σωματιδίων.

### 2.5.1 Σημείο Εξάπλωσης και η Μετάβαση του Μάρκοβ

Για λόγους απλότητας, ας υποθέσουμε πρώτα ότι ο ιδανικός τομέας εικόνας είναι  $R^2$  έτσι ώστε να μην προκύψει καμία αχρείαστη οριακή τεχνικότητα.

Μια δυσδιάστατη λειτουργία σημείου εξάπλωσης (ΛΣΕ)  $K(x, y)$  είναι μια ομαλή λειτουργία που ικανοποιεί

$$(α) \int_{R^2} K(x, y) = 1, \quad \hat{K}(0, 0) = 1; \\ \text{ή ισοδύναμα}$$

$$(β) \hat{K}(\omega_1, \omega_2) \text{ φθίνει αρκετά γρήγορα όπως } |\omega| = \sqrt{\omega_1^2 + \omega_2^2} \rightarrow \infty.$$

Εδώ το  $\hat{K}$  αντιπροσωπεύει τον δυσδιάστατο μετασχηματισμό του Φουριέ

$$\hat{K}(\omega) = \hat{K}(\omega_1, \omega_2) = \int_{R^2} K(x, y) e^{-i(\omega_1 x + \omega_2 y)} dx dy.$$

Στην επεξεργασία σημάτων, αποκαλούνται οι συνθήκες *βαθυπερατού*, εφόσον σε συνδυασμό όταν εφαρμόζονται στις εικόνες, τέτοιες ΛΣΕ διατηρούν σχεδόν συστατικά χαμηλής συχνότητας (συνθήκη (α)) ενώ πιέζει αυτά της υψηλής συχνότητας (συνθήκη (β)).

Μια *ακτινικά* συμμετρική ΛΣΕ  $K$  πιστεύεται ότι είναι *ισοτροπική*. Με άλλα λόγια υπάρχει μια μοναδική μεταβλητή λειτουργία  $k(r^2)$  έτσι ώστε

$$K(x, y) = k(x^2 + y^2).$$

Πιο γενικά το  $K$  πιστεύεται ότι είναι προσανατολιστικά επιλεκτικό ή πολωμένο αν

$$K(x, y) = k((x, y)A(x, y)^T) = k(ax^2 + 2bxy + cy^2),$$

για κάποιους θετικούς οριστικούς πίνακες  $A = [a, b; b, c]$ .

Μια μη-αρνητική PSF  $K(x, y) \geq 0$  για  $(x, y) \in \mathbb{R}^2$  θα μπορούσε να αντιμετωπιστεί φυσικά σαν μια λειτουργία πυκνότητας πιθανότητας λόγω της συνθήκης (α), η οποία είναι επίσης ίση στη συνθήκη του Μάρκοβ για τις μεταβάσεις του Μάρκοβ όπως θα γίνει ξεκάθαρο πιο κάτω.

Δεδομένης μιας PSF  $K(x, y)$ , η συνδεδεμένη εξάπλωση ή ο μετασχηματισμός θολώματος για εικόνες δίνεται από τη συνέλιξη

$$v(x, y) = K * u(x, y) = \int_{\mathbb{R}^2} K(x-p, y-q)u(p, q) dpdq. \quad (2.34)$$

Στον

τομέα του Φουριέ, είναι απλά  $\hat{v} = \hat{K} \cdot \hat{u}$ . Συγκεκριμένα,  $\hat{v}(0, 0) = \hat{u}(0, 0)$ , το οποίο είναι ακριβώς ο νόμος συντήρησης αν  $u(x, y)$  είναι κατανοητό σαν μια λειτουργία πυκνότητας φωτονίου για οπτικά σήματα:

$$\int_{\mathbb{R}^2} v(x, y) dx dy = \int_{\mathbb{R}^2} u(x, y) dx dy.$$

Η συμπεριφορά εξάπλωσης θα μπορούσε να οπτικοποιηθεί εύκολα στην ιδανική περίπτωση όταν

$u = 1_R(x, y)$  και  $K = 1_r(x, y)/(\pi r^2)$ , ο δείκτης λειτουργιών των δίσκων

$B_R$  και  $B_r$  επικεντρωμένο στο  $(0, 0)$  με ακτίνες  $R$  και  $r$ . Τότε μπορεί να αποδειχθεί εύκολα ότι

το  $v = K * u$  υποστηρίζεται πάνω από το  $B_{R+r}$  και εξατμίζεται εκτός. Με άλλα λόγια, η

πληροφορία στην αρχική εικόνα στο  $B_R$  εξαπλώνεται σε μια διευρυμένη περιοχή  $B_{R+r}$ .

Όταν το  $K$  είναι μη αρνητικό, το σημείο εξάπλωσης μεταφέρει την όμορφη στοχαστική ερμηνεία από

τις μεταβάσεις του Μάρκοβ ή τις τυχαίες βόλτες. Ας υποθέσουμε ότι το  $u(x, y)$  υποδηλώνει την χωρική πυκνότητα των φωτονίων. Φανταστείτε (το οποίο μπορεί να μην είναι φυσικά εφικτό!) ότι το κάθε φωτόνιο κινείται τυχαία στο οπτικό πεδίο  $(x, y)$  και υπόκειται στο νόμο μονοβηματικής μετάβασης



$$P(x, y|p, q)dxdy = K(x - p, y - q)dxdy, \quad (2.35)$$

ο οποίος περιγράφει το ποσοστό των φωτονίων στο απειροελάχιστο κουτί  $(p, p + dp) \times (q, q + dq)$  για να μελετηθεί εσωτερικά  $(x, x + dx) \times (y, y + dy)$  μετά από ένα βήμα. Η συνθήκη του Μάρκοβ για τυχαίους βηματισμούς ικανοποιείται αυτόματα απο:

$$\int_{R^2} P(x, y|p, q) dxdy = 1$$

για οποιοδήποτε  $(p, q)$ , το οποίο απλά λέει ότι κανένα φωτόνιο δεν δημιουργείται ή εκμηδενίζεται και κάθε φωτόνιο είτε παραμένει στον αρχικό χώρο ή μεταφέρεται σε νέα τοποθεσία μετά από ένα βήμα. Αν υποθέσουμε ότι η αρχική διανομή φωτονίων δίνεται από  $v(x, y)$  και δηλώνει τη νέα διανομή μετά από ένα βήμα από το  $v(x, y)$ . Με βάση το νόμο μετάβασης, από τα φωτόνια  $u(p, q)dpdq$  το ποσοστό  $P(x, y|p, q)dxdy$  θα μεταφερόταν στο  $(x, x + dx) \times (y, y + dy)$ . Συνεπώς,

$$v(x, y) dxdy = dxdy \int_{R^2} P(x, y|p, q)u(p, q) dpdq, \quad (2.36)$$

το οποίο είναι ακριβώς ο μετασχηματισμός εξάπλωσης του  $u$  όταν το  $P$  δίνεται από το (2.35).

Προσέξτε ότι μέχρι στιγμής ο ακριβής χρόνος αναμονής  $\delta t$  του μοναδικού μεταβατηκού βήματος έχει αγνοηθεί. Στο συνεχόμενο και επαναστατικό σκηνικό, η μετάβαση πρέπει να είναι εξαρτώμενη από το χρόνο:

$$P = P_t(x, y|p, q), \quad t \geq 0.$$

Για την Μπραουνιακή κίνηση [164, 228] για παράδειγμα, κάποιος έχει  $P_{t+s} = P_t \circ P_s$ , όταν το  $P_t$  είναι κατανοητό ως τελεστές μετασχηματισμού πυκνότητας όπως στο (2.36). Επιπλέον, ο νόμος της δυναμικής μετάβασης για την Μπραουνιακή κίνηση είναι πολύ γνωστό ότι είναι

$$P_t(x, y|p, q) = \frac{1}{2\pi t} e^{-\frac{(x-p)^2 + (y-q)^2}{2t}},$$

η δυσδιάστατη ιστροπική Γκαουσιανή με μεταβλητή  $2t$  (ένα  $t$  για κάθε διάσταση).

Πιο γενικά στην ανάλυση εικόνας και όρασης, η σημειακή εκάπλωση μέσω τυχαίων βηματισμών θα μπορούσε να επιτευχθεί από τους συνδεδεμένους απειροελάχιστους παραγωγούς της παραβολικής ΛΣΕ, την οποία θα συζητήσουμε τώρα.

## 2.5.2 Γραμμικό Φιλτράρισμα και Διάχυση

Ας αρχίσουμε με το ψηφιακό σκηνικό του οποίου το  $(u_{ij})$  είναι μια δεδομένη ψηφιακή εικόνα που καθορίζεται από το κανονικό Καρτεσιανό πλέγμα  $(i, j) \in \mathbb{Z}^2$ . Στη θεωρία γραμμικού φιλτραρίσματος, κάποιος έχει

$$\tilde{u}_{ij} = u_{ij} + \varepsilon \sum_{(k,l) \sim (i,j)} h_{ij,kl} u_{kl},$$

όπου το  $\varepsilon > 0$  είναι ένα σταθερό βάρος, και τα  $h$ 's είναι συντελεστές φίλτρου. Η εγγραφή

$(k, l) \sim (i, j)$  αναφέρεται σε όλα τα μικροστοιχεία εικόνας  $(k, l)$  που «συνδέονται» στο  $(i, j)$ . Η συνδετικότητα συνήθως επιτυγχάνεται καθορίζοντας τα γειτονικά παράθυρα, ή άκρες στη θεωρία γραφημάτων [87]. Για να σεβαστούμε το νόμο της συντήρησης, το φίλτρο  $h$  πρέπει να είναι υψιπερατό με την έννοια του

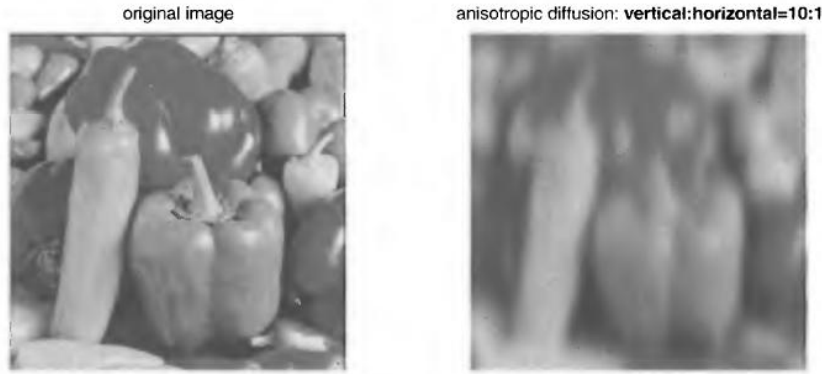
$$\sum_{(k,l) \sim (i,j)} h_{ij,kl} = 0 \quad \text{at each } (i, j).$$

Θεωρείστε, για παράδειγμα, μια μεμβράνη με 5 ιχνοστοιχεία

$$(k, l) \sim (i, j) \quad \text{if and only if} \quad |k - i| + |l - j| \leq 1 \quad (2.37) \quad \text{και το}$$

συσχετισμένο Λαπλασιανό φίλτρο

$$h_{ij,kl} = 1 \text{ if } |k - i| + |l - j| = 1; \quad \text{and} \quad h_{ij,ij} = -4.$$



**Εικόνα**

2.11. Ένα παράδειγμα γραμμικής ανισοτροπικής διάχυσης από το (2.38) με διαγωνικό πίνακα  $D = \text{diag}(D_x, D_y)$  Συνεπώς η εικόνα διαχύεται πολύ πιο γρήγορα κατα μήκος της οριζόντιας κατεύθυνσης  $y$ .

Τότε κάποιος

$$\tilde{u}_{ij} - u_{ij} = \varepsilon [(u_{i+1,j} + u_{i-1,j} - 2u_{ij}) + (u_{i,j+1} + u_{i,j-1} - 2u_{ij})].$$

έχει

Εισάγοντας το χωρικό βήμα  $\Delta x = \Delta y$  και το χρονικό βήμα  $\Delta t$ , επιλέγουμε

$$\varepsilon = D\Delta t / (\Delta x)^2$$

με τέτοιο τρόπο που το  $D$  είναι μια δεδομένη σταθερά σειράς 1.

Τότε

$$\frac{\tilde{u}_{ij} - u_{ij}}{\Delta t} = D \left( \frac{u_{i+1,j} + u_{i-1,j} - 2u_{ij}}{(\Delta x)^2} + \frac{u_{i,j+1} + u_{i,j-1} - 2u_{ij}}{(\Delta y)^2} \right).$$

έστω ότι  $\Delta t$  και το  $\Delta x = \Delta y$  περνούν προς το 0, κάποιος καταλήγει με τη σταθερή εξίσωση θέρμης

$$\frac{\partial u}{\partial t} = D \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) = D\nabla^2 u,$$

δεδομένου ότι  $u_{ij} = u(i\Delta x, j\Delta y, t)$  και

$$\tilde{u}_{ij} = u(i\Delta x, j\Delta y, t + \Delta t).$$

Έτσι η επανάληψη αυτού του κλασσικού γραμμικού φίλτρου είναι ισάξιο της επεξεργασίας διάχυσης.

Πιο γενικά επιτρέποντας στο  $D = D(x, y)$  να διαφέρει χωρικά ή έστω να είναι ένας

θετικός οριστικός πίνακας 2 επί 2  $D = [a(x, y), b(x, y); b(x, y), c(x, y)]$ ,

κάποιος πέρνει την έκδοση ΛΣΕ των χωρικά διαφοροποιημένων γραμμικών στάσιμων φίλτρων:

$$u_t = \nabla \cdot [D\nabla u] = (au_x)_x + (bu_y)_x + (bu_x)_y + (cu_y)_y. \quad (2.38)$$

Αυτές οι εξισώσεις συνοδεύονται από κατάλληλες οριακές συνθήκες καθώς και την αρχική συνθήκη  $u(x, y, 0) = u_0(x, y)$  τη δεδομένη εικόνα. Η Εικόνα 2.11 δείχνει ένα παράδειγμα.

Φυσικά, η διάχυση είναι μια τυπικά αυτο-οδηγούμενη επαναστατική διαδικασία και η κινητήριος «δύναμη» είναι χωρικά ανομοιογενής. Είναι επίσης έκδηλη στην φόρμουλα ροής

$$j = -D \nabla u = -D(u_x, u_y)^T.$$

Εφόσον το  $D = D(x, y) = [a, b; b, d]$  είναι θετικά οριστικό, τέτοια ροή πρέπει να κλίνει καθοδικά εφόσον

$$\frac{\partial u}{\partial t} = j \cdot \nabla u = -(\nabla u)^T D (\nabla u) < 0.$$

Συνεπώς η διάχυση αποδυναμώνει την ανάπτυξη μεγάλων κλίσεων και εξαλύπτει γρήγορες τοπικές ταλαντώσεις που προκαλούνται από τυχαίο θόρυβο.

Επιπλέον, κατά τη διάρκεια μιας επεξεργασίας γραμμικής διάχυσης, ταλαντώσεις με διαφορετικούς χωρικούς κυματικούς αριθμούς μειώνονται με διαφορετικούς ρυθμούς. Για απλότητα, πρώτα θεωρείστε την περίπτωση όταν  $\Omega = \mathbb{R}^2$  και  $D = [a, b; b, c]$  είναι ένας θετικός οριστικός σταθερός πίνακας. Υποδηλώνεται από το  $L$  ο χωρικός γραμμικός τελεστής. Μια μονοχρωματική ταλάντωση στην κατεύθυνση του  $\theta$  με χωρική συχνότητα  $k$  δίνεται από

το  $\phi(x) = e^{i(k \cdot x)}$  με  $k = (k \cos \theta, k \sin \theta)$  και  $x = (x, y)$ . Τότε

$$L\phi = -k D k^T e^{i(k \cdot x)} = -\lambda \phi, \quad (2.39)$$

με  $\lambda = k D k^T > 0$ . Έτσι το  $\phi$  είναι αναλλοίωτο στο  $L$  εφόσον το  $L\phi$  βρίσκεται ακόμη στην «κατεύθυνση» του  $\phi$ , ή γενικά μιλώντας το  $\phi$  είναι το αυτοδιάνυσμα του  $L$  συνδεδεμένο με την ιδιοτιμή  $\lambda$ . Η μη-διακύμανση επιτρέπει σε κάποιον να περιορίζει την επεξεργασία διάχυσης  $u_t = Lu$  εντός του διαστήματος  $\phi$  και να προσπαθεί μια λύση υπό τη μορφή  $u(x, t) = g(t)\phi(x)$ . Τότε

$$g'(t) = -\lambda g(t) \quad \text{and} \quad g(t) = g(0)e^{-\lambda t}, \quad (2.40)$$

το οποίο δείχνει ότι ο ρυθμός μείωσης για το  $\phi = e^{i(k \cdot x)}$  είναι  $\lambda = k D k^T$ . Εφόσον το  $D$  είναι ένας σταθερός θετικός οριστικός πίνακας, κάποιος έχει  $\lambda = O(k^2)$ ,

επιβεβαιώνοντας την προηγούμενη υπόθεση ότι οι γρηγορότερες ταλαντώσεις πεθαίνουν πιο γρήγορα κατά τη διάρκεια της επεξεργασίας διάχυσης.

Η πιο πάνω ανάλυση μεταφέρεται όμορφα στις επεξεργασίες γραμμικής διάχυσης σε γενικούς τομείς περιορισμένης εικόνας με κατάλληλες οριακές συνθήκες. Το εγναρκτήριο σημείο είναι να λύσουμε

πρώτα τις αναλλοίωτες εξισώσεις (2.39):  $L\phi = \lambda\phi$  για κατάλληλα  $\lambda$ 's που θα χαρακτηρίσουν διαδορτικούς ρυθμούς ταλαντώσεων στον τομέα. Είναι ένα πρόβλημα ιδιοτιμής εφόσον το  $L$ , μαζί με τις οριακές συνθήκες συμπεριφέρεται κατά πολύ σαν ένας μη-αρνητικός πίνακας. Οι τετραγωνισμένες συνθήκες είναι γενικά κβαντοποιημένες στο

$$0 \leq \lambda_1 \leq \lambda_2 \leq \dots,$$

και οι συνδεδεμένες ιδιοσυναρτήσεις απεικονίζουν τα συνδεδεμένα χωρικά μοτίβα. Τότε εξακολουθεί να ισχύει ο νόμος της διάσπασης (2.40).

Ακολουθώντας αυτή τη γραμμή, κάποιος παρατηρεί το έντονο χαρακτηριστικό της διάχυσης: όπως

$t \rightarrow \infty$ , όλοι οι συντελεστές ταλάντωσης με μη-μηδενικά  $\lambda$ 's πρέπει να εξαφανιστούν. Κάτω από την αδιαβατική οριακή συνθήκη του Νούμαν για παράδειγμα, μόνος ο άμεσος-παρόν συντελεστής

$\phi_1 \equiv 1$ : συνδεδεμένος με  $\lambda_1 = 0$  παραμένει στο τέλος. Επιπλέον, από το νόμο συντήρησης (εγγυημένο από την αδιαβατική συνθήκη), κάποιος πρέπει να έχει

$$u(x, t) \rightarrow c_1 \phi_1 = \frac{1}{\text{area}(\Omega)} \int_{\Omega} u_0(x) dx \quad \text{as } t \rightarrow \infty.$$

Είναι

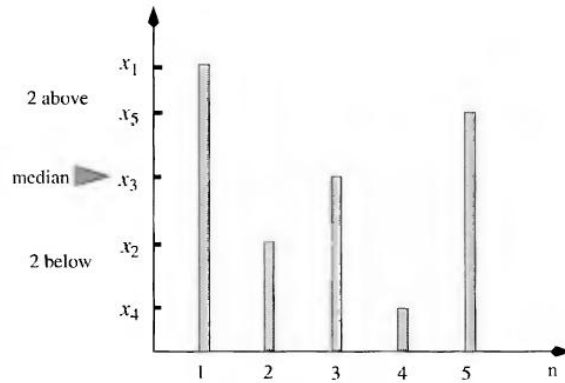
βέβαια ανεπιθύμητο στην επεξεργασία εικόνας εφόσον όλα τα χαρακτηριστικά εικόνας διαγράφονται.

Συνεπώς, τόσο στην πράξη όσο και στη θεωρία, είναι ένα σοβαρό πρόβλημα να αποφασίσουμε πότε θα σταματήσουμε την επεξεργασία διάχυσης προτού όλα τα κύρια χαρακτηριστικά εικόνας και τα

οπτικά μοτίβα έχουν εξαφανιστεί. Αυτή τη φορά το  $T_s$  αποκαλείται ο *άριστος χρόνος σταματήματος*. Η απάντηση εξαρτάται ξεκάθαρα από την εκάστοτε διεργασία. Για κάθε εξάλειψη λευκού θορύβου με

μεταβλητή  $\sigma^2$  για παράδειγμα θα μπορούσε να αποδειχτεί ότι  $T_s = O(\sigma^2)$ . Για χωρικά ανομοιογενείς θορύβους, κάποιος θα μπορούσε να φανταστεί ότι ένας ομοιόμορφος χρόνος

σταματήματος  $T_s$  δεν μπορεί να αρκεί για να χαρακτηρίσει πλήρως την χωρική ανομοιογένεια, εκτός αν ο συντελεστής διάχυσης  $D$  έχει ήδη προσαρμοστεί στο θόρυβο.



Εικόνα

$(x_n)$ .

2.12 Ο μέσος όρος ενός τμήματος 5 συντελεστών από ένα φανταστικό μονοδιάστατο σήμα

### 2.5.3 Μη-γραμμικό Φιλτράρισμα και Διάχυση

Σε σύγκριση με τα πλείστα ακουστικά ή ηχητικά σήματα, οι εικόνες διαφέρουν στο ότι είναι μη συνεχόμενες λειτουργίες. Οι ασυνέχειες στις δυσδιάστατες εικόνες είναι ευρέως συνδεδεμένες με τα όρια των αντικειμένων στον τρισδιάστατο κόσμο και συνεπώς ενδογενής και οπτικά σημαντικές. Το αποτελεσματικό φιλτράρισμα και οι επεξεργασίες διάχυσης πρέπει να μπορούν να ξεχωρίσουν τέτοιες ιδιαιτερότητες από θορυβώδης ταλαντώσεις. Οι γραμμικές διαχύσεις αποτυγχάνουν από αυτή την άποψη λόγω της έλλειψης προσαρμογής σε δεδομένες εικόνες. Έτσι η μη-γραμμικότητα είναι φυσικά απαιτούμενη στην ανάλυση και επεξεργασία εικόνων.

Το πιο γνωστό και απλούστερο μη-γραμμικό φίλτρο είναι ίσως το μέσο φίλτρο:

$$\tilde{u}_{ij} = \text{median}\{u_{kl} : (k, l) \sim (i, j)\}. \quad (2.41)$$

Εδώ

όπως και στις προηγούμενες υποενότητες, η συνδεσιμότητα ανάμεσα σε μικροστοιχεία εικόνας  $\sim$  ορίζεται συχνά από ένα τοπικό παράθυρο, για παράδειγμα, η Καρτεσιανή γειτονιά 5 επί 5. Η Εικόνα 2.12 δείχνει ένα μονοδιάστατο παράδειγμα με ένα παράθυρο μεγέθους 5. Το μέσο φιλτράρισμα, είναι ένα ιδιαίτερο παράδειγμα της πιο γενικής τάξης τριτοβάθμιας φιλτραρίσματος στατιστικής

τάξεως. Ας ανακαλέσουμε ότι για τις παρατηρήσεις  $2n + 1$

$(x_{-n}, \dots, x_0, \dots, x_n)$  (ένα παράθυρο στο περιθώριο  $(\cdot)$ ), οι στατιστικές τάξεως

$x_{(-n)} \leq x_{(-n+1)} \leq \dots \leq x_{(n)}$  έχουν εδραιωθεί ταξινομώντας τα δεδομένα με αύξουσα σειρά. Τότε ένα γενικό βαθυπερατό φίλτρο στατιστικής τάξεως  $H$  καθορίζεται από

$$\tilde{x}_0 = H(x_{-n}, \dots, x_n) = h_{-n}x_{(-n)} + \dots + h_nx_{(n)} \quad (2.42)$$

για

κάποιο κατάλληλο σύνολο συντελεστών φίλτρων  $h$ 's. Προσέξτε ότι η μη γραμμικότητα προκαλείται από τη σύνδεση των σταθερών συντελεστών φίλτρου με τις στατιστικές τάξεως, αντί με τα αρχικά ωμά

δεδομένα. (Η επιλογή του  $\tilde{x}_0$  είναι καθαρά για χάρη απόδειξης και κατά κύριο λόγο κάποιος θα

μπορούσε χρησιμοποιήσει την ίδια έκφραση για να αναβαθμίσει την αξία σε οποιοδήποτε σταθερό περιθώριο  $m$  ανάμεσα σε  $-n$  και  $n$ .) Για το μέσο φιλτράρισμα,  $h_k = \delta_k$ , η διακριτική ακολουθία δέλτα του Ντιράκ.

Λόγω της στατιστικής φύσης (δηλαδή μέσα σε ένα παράθυρο με αρκετά δείγματα δεδομένων), το μέσο φίλτρο εφαρμόζεται εύκολα για την αφαίρεση θορύβου. Αλλά γιατί είναι ικανό να διατηρεί τις άκρες; Σαν παράδειγμα ας θεωρήσουμε ένα απλό μονοδιάστατο «κρήμνο» ή ένα σήμα τύπου Χέβισαϊτ:

$$x = (\dots, 1.4, 1.3, 1.2, 1.1, 1, 0, 0.1, 0.2, 0.3, 0.4, \dots).$$

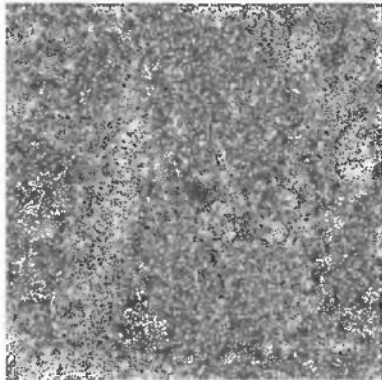
Ας

υποθέσουμε ότι  $x_0 = 1$ , όπου βρίσκεται ο κρήμνος. Εφαρμόζοντας το μέσο φίλτρο με ένα συμμετρικό παράθυρο μήκους 5 (δηλαδή (i.e.,  $n = 2$ ),

κάποιος εξασφαλίζει την εξαγόμενη ακολουθία κοντά στον κρήμνο,

$$\tilde{x}_{-2} = 1.2, \tilde{x}_{-1} = 1.1, \tilde{x}_0 = 1, \tilde{x}_1 = 0.2, \tilde{x}_2 = 0.2.$$

salt & pepper noise with 40% spatial density



a single pass of median filtering with a 7x7 window



**Εικόνα 2.13.** Ένα παράδειγμα χρήσης μέσου φίλτρου (με επικεντρωμένο τετράγωνο παράθυρο 7 x 7) για την αποθορύβωση μιας εικόνας με έντονο θόρυβο από αλάτι και πιπέρι. Προσέξτε το ξεχωριστό χαρακτηριστικό του μέσου φίλτρου: οι άκρες στην αποκατεστημένη εικόνα δεν είναι θολές.

Παρόλο που οι αξίες στο  $m = 1$  και 2 φαίνονται να είναι ενημερωμένες ασαφώς για ένα τέτοιο τμηματικά ομαλό σήμα, ο απότομος κρήμνος στο  $m = 0$  είναι πράγματι καλά συντηρημένος. Η Εικόνα 2.13 δείχνει ένα παράδειγμα αποθορύβωσης εικόνας μέσω μέσου φιλτραρίσματος.

Τώρα θεωρήστε ένα γενικό ιδανικό δισδιάστατο κρήμνο Χέβισαϊτ καθορισμένο από

$$u(x, y) = H(ax + by + c), \quad (x, y) \in \mathbb{R}^2, \quad \text{όπου}$$

$a, b, c$  είναι σταθερές και  $H = H(s)$  είναι η λειτουργία Χέβισαίτ:

$H(s) = 1, s \geq 0$ , και αλλιώς  $H(s) = 0$ . Ο κρήνος εξαπλώνεται στην ευθεία

γραμμή  $ax + by + c = 0$ . Για οποιοδήποτε μικροστοιχείο εικόνας  $(x_+, y_+)$  έτσι

ώστε  $ax_+ + by_+ + c > 0$ , οποιοδήποτε μέσο φίλτρο με ένα ακτινικά συμμετρικό παράθυρο (δηλαδή δίσκο) θα οδηγούσε στο

$$\tilde{u}(x_+, y_+) = 1 = u(x_+, y_+),$$

απλά επειδή πιο πολλά μικροστοιχεία μέσω στο δίσκο παραθύρου έχουν την αξία 1 παρά 0. Όμοια,

για οποιοδήποτε μικροστοιχείο εικόνας  $(x_-, y_-)$  με

$$ax_- + by_- + c < 0,$$

$$\tilde{u}(x_-, y_-) = 0 = u(x_-, y_-).$$

Συνεπώς, το αποτέλεσμα από το μέσο φιλτράρισμα συντηρεί στον κρήμο στην υψηλά ιδανικοποιημένη του περίπτωση.

Το μη-γραμμικό φιλτράρισμα έχει γίνει ένα πολύ σημαντικό θέμα προς μελέτη στο ψηφιακό σήμα και την επεξεργασία εικόνας [63]. Τώρα στρεφόμαστε στην τάξη μη-γραμμικών ψηφιακών φίλτρων τα οποία είναι συνδεδεμένα με μη-γραμμικές διαχύσεις.

Προκειμένου να συντηρήσουμε μακροσκοπικά τις άκρες, τα τυχαία κινούμενα σωματίδια δεν πρέπει να περάσουν τις άκρες των «κρημών». Έτσι στη φόρμουλα μετάβασης πυκνότητας

$$\tilde{u}_{ij} = \sum_{(k,l) \sim (i,j)} P_{ij,kl}^\varepsilon u_{kl},$$

οι πιθανότητες μετάβασης  $P_{ij,kl}^\varepsilon$ 's πρέπει να είναι προσαρμόσιμες. Για παράδειγμα, θεωρώντας

το  $\mathbf{u}$  ως μια λειτουργία ενέργειας, όπως παρακινείται από το νόμο διανομής του Γκίμπς στην προηγούμενη ενότητα, κάποιος μπορεί να ορίσει



$$P_{ij,kl}^\varepsilon = P_{ij,kl}^\varepsilon(u) = \frac{1}{Z^\varepsilon} e^{-\frac{|u_{ij}-u_{kl}|^\alpha}{\varepsilon}}, \quad (k, l) \sim (i, j). \quad (2.43)$$

Εδώ η ελεγχόμενη παράμετρος  $\varepsilon$  παίζει το ρόλο της θερμοκρασίας, και η λειτουργία σωματιδίων  $Z^\varepsilon$  κανονικοποιεί τη συνολική εξερχόμενη πιθανότητα μετάβασης στο 1:

$$\sum_{(i,j) \sim (k,l)} P_{ij,kl}^\varepsilon = 1 \quad \text{for any } (k, l). \quad (2.44)$$

Η δυναμική παράμετρος  $\alpha$  είναι επίσης συντονίσιμη. Με αυτό το νόμο, οι μεταβάσεις ανάμεσα σε δυο γειτονικά μικροστοιχεία με μεγάλη διαφορά στο  $u$  είναι πολύ απίθανες.

Για να αποσπάσουμε τις μακροσκοπικές διαχύσεις ΛΣΕ, ας υποθέσουμε ότι το

$u_{ij} = u(ih, jh)$  με το  $\Delta x = \Delta y = h$  και ότι ισχύει η συνθήκη βαθυπερατού:

$$\sum_{(k,l) \sim (i,j)} P_{ij,kl}^\varepsilon = 1 \quad \text{at any } (i, j). \quad (2.45)$$

(Με το  $P$  να ικανοποιεί τόσο το (2.44) όσο και το (2.45) πιστεύεται ότι είναι διπλά στοχαστικό.) Τότε η πιο πάνω φόρμουλα μετάβασης πυκνότητας γίνεται:

$$\tilde{u}_{ij} - u_{ij} = \sum_{(k,l) \sim (i,j)} P_{ij,kl}^\varepsilon (u_{kl} - u_{ij}).$$

Ας υποθέσουμε ότι η γειτονική συνδεσιμότητα καθορίζεται από μεμβράνες 5 μικροστοιχείων όπως στη γραμμική περίπτωση προηγουμένως, και οι πιθανότητες μετάβασης φέρουν τη μη-γραμμική μορφή του

$$P_{ij,kl}^\varepsilon = P^\varepsilon \left( u_{\frac{i+k}{2}, \frac{j+l}{2}}, \nabla u_{\frac{i+k}{2}, \frac{j+l}{2}} \right) \quad (2.46)$$

για

κάποια κατάλληλη συνεχόμενη λειτουργία  $P^\varepsilon(u, \nabla u)$ . Για να είμαστε ακριβείς, τώρα

υιοθετούμε την εγγραφή του  $P_{\frac{i+k}{2}, \frac{j+l}{2}}^\varepsilon$  για τέτοιο  $P_{ij,kl}^\varepsilon$ . Τότε

$$\begin{aligned}\tilde{u}_{ij} - u_{ij} &= P_{i+\frac{1}{2},j}^\varepsilon (u_{i+1,j} - u_{ij}) - P_{i-\frac{1}{2},j}^\varepsilon (u_{ij} - u_{i-1,j}) + \dots \\ &\simeq h \left[ (P^\varepsilon \partial_x u)_{i+\frac{1}{2},j} - (P^\varepsilon \partial_x u)_{i-\frac{1}{2},j} \right] + \dots \\ &\simeq h^2 \partial_x (P^\varepsilon \partial_x u) + h^2 \partial_y (P^\varepsilon \partial_y u).\end{aligned}$$

Εδώ οι δυο διαφορετικοί υπολογισμοί διαφοροποίησης είναι ακριβείς στις ταξικές οδηγίες στο

$h$ . Τώρα επιλέγουμε  $P^\varepsilon$  έτσι ώστε

$$P^\varepsilon = P^\varepsilon(u, \nabla u) = \frac{\varepsilon}{h^2} D(u, \nabla u)$$

για κάποια σταθερή θετική λειτουργία  $D$ .

Αναγνωρίζοντας  $u_{ij} = u(ih, jh, t)$  και

$$\tilde{u}_{ij} = u(ih, jh, t + \varepsilon),$$

συνεπώς εξασφαλίζουμε τη μη-γραμμική εξίσωση διάχυσης στο όριο  $\varepsilon, h \rightarrow 0$ :

$$u_t = \nabla \cdot D(u, \nabla u) \nabla u.$$

Προσέξτε ότι για να γίνει έγγυρη η συνθήκη βαθυπερατού (2.45), αρκεί να απαιτήσουμε

$$P^\varepsilon \leq 1/4, \quad \varepsilon \|D\|_\infty / h^2 \leq 1/4.$$

Στην υπολογιστική ΣΛΕ, αυτό είναι πού γνωστό ως η συνθήκη CFL που ονομάστηκε χάρη στους Κουράντ (Courant), Φράντριχς (Friedrichs) και Λιούι (Lewy) [287]. Στο [251], οι Περόνα και Μάλικ πρότειναν για πρώτη φορά να επιλέξουμε

κατάλληλα την εξάρτηση του  $D$  στο  $\nabla u$  προκειμένου να συντηρηθούν οι άκρες στις εικόνες.

## Κεφάλαιο 3

# Μοντελοποίηση και Αναπαράσταση Εικόνων

Ο λόγος για τη μοντελοποίηση ή την αναπαράσταση εικόνων είναι να βρεθούν κατάλληλοι τρόποι για να περιγράψουμε και να αναλύσουμε μαθηματικά εικόνες. Παρόλα αυτά, κάποιος πρέπει να αντιληφθεί ότι δεν υπάρχει απόλυτα η καλύτερη αναπαράσταση, εφόσον η τελειότητα εξαρτάται αναπόφευκτα από συγκεκριμένες διεργασίες επεξεργασίας, όπως στην περίπτωση διαφορετικών αναπαραστάσεων φυσικών αριθμών: το δεκαδικό σύστημα είναι πιο βολικό από το δυαδικό στην καθημερινή ζωή, αλλά το τελευταίο είναι πιο φυσικό για ψηφιακούς ή κβαντικούς υπολογιστές. Το παρόν κεφάλαιο εισάγει πέντε γενικές και χρήσιμες προσεγγίσεις για την αναπαράσταση εικόνων, με βάση τις οποίες θα αναπτυχθούν σε μετέπειτα κεφάλαια αρκετοί πετυχημένοι επεξεργαστές εικόνας.

### **3.1 Μοντελοποίηση και Αναπαράσταση: Τι, Γιατί και Πώς**

Οι ψηφιακές εικόνες παρουσιάζονται ως επί των πλείστων σαν ένας πίνακας κλιμακώσεων για εικόνες γκριζας κλίμακας ή φορείς για έγχρωμες εικόνες, εφόσον συχνά απαντιούνται από συστοιχίες διάταξης σύζευξης φορτίου (Charge Coupled Device - CCD) ή παρουσιάζονται από συστοιχίες υγρών κρυστάλλων (όπως για τους φορητούς υπολογιστές ή για τους υπολογιστές τσέπης). Η αναπαράσταση πίνακα μικροστοιχείων εικόνας δεν είναι, παρόλα αυτά, σε καμιά περίπτωση πιο επαρκής από την άποψη πληροφορίας – θεωρίας.

Μια αναπαράσταση μιας δεδομένης τάξης  $U$  φυσικών εικόνων αναφέρεται στο μετασχηματισμό

$u = (u_{i,j})$  κάτω από τον οποίο οποιαδήποτε εικόνα

$$u = u(x, y) \text{ with } (x, y) \in \Omega = (a, b) \times$$

της τάξης μετασχηματίζεται σε ένα νέο τύπο δεδομένων ή δομή  $w = \mathcal{T}u$ . Έστω ότι το

$W$  υποδηλώνει χωρική ευρύτητα του  $\mathcal{T}$ , ο οποίος καλείται επίσης και διάστημα μετασχηματισμού. Τότε έχουμε, για συντομία,

$$\mathcal{T} : U \rightarrow W, \quad u \rightarrow w = \mathcal{T}u.$$

Μια αναπαράσταση λέγεται ότι είναι γραμμική αν

1. τόσο η ταξική εικόνα  $U$  όσο και το μετασχηματιστικό του διάστημα είναι γραμμικά διαστήματα,

ή κυρτοί κώνοι γραμμικών διαστημάτων, έτσι ώστε για παράδειγμα, το  $au_1 + bu_2$  ανήκει στο  $U$  με την προϋπόθεση ότι το  $u_1$  και το  $u_2$  κάνουν και  $a, b \in \mathbb{R}$  ή  $\mathbb{R}^+$ ; και

2. ο μετασχηματισμός είναι γραμμικός:  $\mathcal{T}[au_1 + bu_2] = a\mathcal{T}u_1 + b\mathcal{T}u_2$ .

Για παράδειγμα, ας θεωρήσουμε την τάξη όλων των εικόνων δυο μικροστοιχείων:

$U = \{u = (s, t) : s, t \in \mathbb{R}^+\}$  η οποία είναι ένας κυρτός κώνος στο  $\mathbb{R}^2$  (δηλαδή το πρώτο τεταρτοκύκλιο). Έστω ότι το διάστημα μετασχηματισμού είναι  $w = (a, d) \in W = \mathbb{R}^2$ , και έστω ότι το  $\mathcal{T}$  είναι η μέση-διαφορά μετασχηματισμού του Χάαρ [96, 290]:

$$w = \mathcal{T}u : a = \frac{s+t}{2} \text{ and } d = \frac{s-t}{2}.$$

Είναι ο γραμμικός μετασχηματισμός κλειδί για τη δόμηση των κυματιδίων του Χάρ [96, 290].

Μαθηματικά, ακόμη και το σύστημα της ανθρώπινης όρασης θα μπορούσε να θεωρηθεί ως ένας (βιολογικός) τελεστής αναπαράστασης  $\mathcal{T}_h$ . Οι φωτοδέκτες (δηλαδή κώνοι και ράβδοι) των δυο αμφιβλίστροειδών είναι πράγματι διαταγμένοι σε μεγάλο βαθμό με ένα δυσδιάστατο τρόπο. Έστω ότι

το  $U$  υποδηλώνει την τάξη των εικόνων  $u = (u_l, u_r)$  που προβάλλονται στους αμφιβλίστροειδείς τόσο του δεξιού όσο και του αριστερού ματιού. Έστω ότι το  $W$  υποδηλώνει την τάξη των ηλεκτροχημικών σημάτων που αποκρυπτογραφούνται από όλους τους κύριους οπτικούς φλοιούς (π.χ.  $V1$  και  $MT$  [227, 168]). Τότε το σύστημα της ανθρώπινης όρασης (από τους φωτοδέκτες, τα γαγγλιακά κύτταρα, και των εξωτερικό γονατώδες πυρήνα (LGN – lateral geniculate nucleus) μέχρι τον κύριο οπτικό φλοιό  $V1$  και λοιπά) είναι ένας βιολογικός μετασχηματισμός  $\mathcal{T}_h$  που δημιουργείται από αυτό το πολύπλοκο δίκτυο νευρών και κυττάρων. Πολυάριθμα στοιχεία δείχνουν ότι είναι μη-γραμμικό [150, 152].

Πώς κρίνει κάποιος εάν ένας συγκεκριμένος τύπος αναπαράστασης εικόνας είναι καλός για την ανάλυση εικόνας και όρασης; Οι εικόνες ως μια σημαντική τάξη σημάτων μεταφέρουν πλούσιες υλικές, γεωμετρικές και τοποθετικές πληροφορίες σχετικά με τον τρισδιάστατο κόσμο. Μια καλή αναπαράσταση θα έπρεπε να μπορεί να δώσει έμφαση σε τέτοιες πληροφορίες και να συλλάβει επιτυχώς τα συνδεδεμένα οπτικά χαρακτηριστικά κλειδί. Αυτή είναι η πιο γενική οδηγία για επιτυχή μοντελοποίηση ή αναπαράσταση εικόνας.

Μια αναπαράσταση  $\mathcal{T}$  λέγεται ότι είναι χωρίς απώλειες αν οποιοδήποτε  $u$  μπορεί αναδομηθεί τέλεια από την αναπαράστασή του  $w = \mathcal{T}u$ . Με άλλα λόγια υπάρχει άλλος ένας (αναδομικός) μετασχηματισμός  $\mathcal{R}$  από το  $W$  στο  $U$  έτσι ώστε  $u = \mathcal{R}[\mathcal{T}[u]]$  για κάθε  $u \in U$ .

Μια απαραίτητη προϋπόθεση για να είναι χωρίς απώλειες μια αναπαράσταση  $\mathcal{T}$  είναι να είναι εγχυτική· δηλαδή, δυο ξεχωριστές εικόνες  $u$  και  $v$  θα έπρεπε να μετασχηματιστούν σε διαφορετικά  $\mathcal{T}u$  και  $\mathcal{T}v$ , εφόσον  $\mathcal{T}u = \mathcal{T}v$  θα υπονοούσε άμεσα ότι

$$u = \mathcal{R}[\mathcal{T}[u]] = \mathcal{R}[\mathcal{T}[v]] = v.$$

Η αναπαράσταση του Χάαρ στο (3.1) είναι προφανώς χωρίς απώλειες εφόσον η αναδόμηση  $\mathcal{R}$  εφαρμόζεται εύκολα από το  $s = a + d$  και  $t = a - d$ .

Πιο γενικά από την άποψη της γραμμικής άλγεβρας, η φύση χωρίς απώλειες της αναπαράστασης του Χάαρ είναι εγγυημένη από την αντιστρεψιμότητα των συνδεδεμένων πινάκων.

Στην ανάλυση σημάτων και εικόνων, οι δυο μετασχηματισμοί  $\mathcal{T}$  και  $\mathcal{R}$  αποκαλούνται συνήθως οι μετασχηματισμοί *ανάλυσης* και *σύνθεσης* ξεχωριστά. Η ανάλυση είναι να αναλύσουμε ένα δεδομένο σήμα εικόνας και να εξάγουμε τα χαρακτηριστικά και πληροφορίες κλειδί, ενώ η σύνθεση είναι η προσπάθεια αναδόμησης του αρχικού σήματος από την εξαγωγή του βήματος ανάλυσης.

Αν κάποιος έχει ένα πλήρες βιβλίο κωδικών ή λεξικό στο χέρι, το είναι εγχυτικός είναι ήδη αρκετό για να είναι χωρίς απώλειες. Για παράδειγμα, αν 10 πολύπλοκες εικόνες είχαν αποθηκευτεί και ονομασθεί

εκ των προτέρων από το  $u_1, u_2, \dots, u_{10}$ , η απλή εγχυτική αναπαράσταση

$$\mathcal{T} : u_k \rightarrow k$$

είναι ήδη αρκετή για να αναδομήσει τις εικόνες εφόσον το βιβλίο κωδικών είναι διαθέσιμο. Αν φανταστούμε ότι ο αριθμός  $k = 3$  έχει ληφθεί, κάποιος μπορεί απλά να στραφεί στο βιβλίο κωδικών για να επανακτήσει την εικόνα  $u_3$ .

Σε αντίθεση με οποιαδήποτε εθνική γλώσσα το λεξιλόγιο της οποίας είναι σχετικά σταθερό και οριστικό, είναι ουσιαστικά αδύνατο να δημιουργήσει ένα λεξικό που να περιέχει όλες τις εικόνες του κόσμου. Ως αποτέλεσμα, η προηγούμενη ιδέα αναπαράστασης βασισμένη στο λεξικό δεν μπορεί να λειτουργήσει αποτελεσματικά στον πραγματικό κόσμο.

Αντιθέτως, οι περισσότερες καλές αναπαραστάσεις επικεντρώνονται στην σύλληψη *ουσιαστικών* οπτικών χαρακτηριστικών εικόνων. Αγνοώντας οπτικά ασήμαντες πληροφορίες εικόνων, μια καλή αναπαράσταση έχει αναπόφευκτα απώλειες αντί να είναι χωρίς απώλειες. Παρόλα αυτά, οι συνδεδεμένες επεξεργασίες αναδόμησης οδηγούν σε λογικά καλές συνθέσεις εφόσον καμία σημαντική οπτική πληροφορία δεν έχει αγνοηθεί. Αυτή είναι πράγματι η περίπτωση στην βασισμένη σε κυματίδιο JPEG2000.

### 3.2 Προσδιοριστικό Μοντέλο Εικόνας

Πρώτα συζητούμε διάφορα προσδιοριστικά μοντέλα εικόνας που επικρατούν στη λογοτεχνία. Είναι όλα μαθηματικά μοντέλα που υπολογίζουν πραγματικές εικόνες σε συγκεκριμένα επίπεδα πιστότητας.

#### 3.2.1 Εικόνες ως Κατανομές (Γενικοποιημένες Λειτουργίες)

Η αντιμετώπιση εικόνων ως κατανομών ή γενικοποιημένων λειτουργιών είναι η πιο ευρεία προσέγγιση για προσδιοριστικά μοντέλα [225] και έχει σημαντικές αξίες στην κατανόηση εικόνων όπως επεξηγείται πιο κάτω.

Όπως και πριν, έστω ότι το  $\Omega$  υποδηλώνει ένα ανοιχτό και περιορισμένο δυοδιάστατο τομέα εικόνας Λιπσχιτζ. Το σύνολο των λειτουργικών ελέγχων προσδιορίζεται ως

$$D(\Omega) = \{\phi \in C^\infty(\Omega) | \text{supp}\phi \subseteq \Omega\}$$

Κάθε λειτουργικός έλεγχος  $\phi \in D(\Omega)$  θα μπορούσε να θεωρηθεί ως ένας γραμμικός ανιχνευτής για την αιχμαλώτηση σημάτων εικόνας.

Μια εικόνα  $u$  στο  $\Omega$ , που αντιμετωπίζεται ως μια κατανομή, είναι μια γραμμική λειτουργικότητα στο  $D(\Omega)$ :

$$u : \phi \rightarrow \langle u, \phi \rangle,$$

όπου το εσωτερικό σύμβολο προϊόντος έχει χρησιμοποιηθεί επίσημα, όπως συνηθίζεται στη κατανομική θεωρία [3, 193]. Συνεπώς, το διάστημα εικόνας είναι το διάστημα κατανομής  $D'(\Omega)$ .

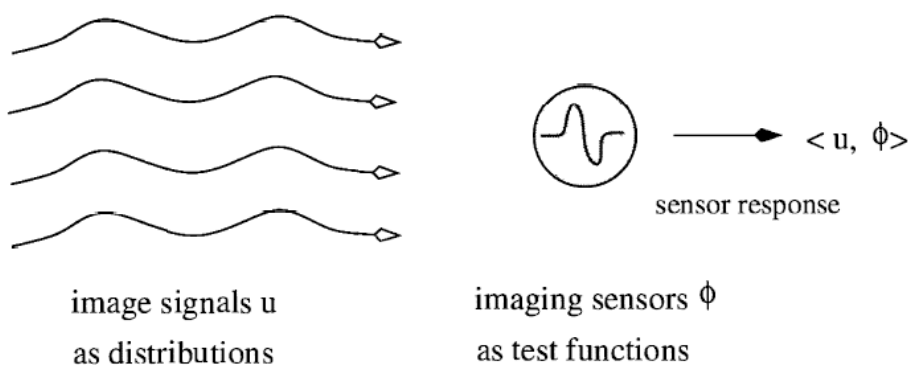
Παρόλο που μια τέτοια εικόνα  $u$  μπορεί να μη μοιάζει καθόλου με οποιαδήποτε γνώριμη λειτουργία, σίγουρα εξάγει μια μοναδική απάντηση  $\langle u, \phi \rangle$  για οποιοδήποτε αισθητήρα  $\phi \in D(\Omega)$ , ο οποίος επιχειρεί να ανιχνεύσει την ύπαρξη και τα χαρακτηριστικά της εικόνας  $u$ . Επιπλέον, η ανίχνευση είναι γραμμική:

$$\langle u, a\phi + b\psi \rangle = a\langle u, \phi \rangle + b\langle u, \psi \rangle \quad \forall \phi, \psi \in D(\Omega), \quad a, b \in \mathbb{R}.$$

Η ιδέα της κατανομής είναι τόσο εύκαμπτη που κάποιος μπορεί να πάρει παράγωγα με οποιαδήποτε σειρά. Στην περίπτωση δυσδιάστασης για παράδειγμα, έστω ότι  $\alpha_1$  και  $\alpha_2$  είναι δυο μη-αρνητικοί ακέραιοι αριθμοί. Τότε για οποιαδήποτε κατανομή  $u \in D'(\Omega)$ , το παράγωγο του  $v = \partial_1^{\alpha_1} \partial_2^{\alpha_2} u$  καθορίζεται ως μια νέα κατανομή (ή γραμμική λειτουργικότητα) έτσι ώστε για οποιοδήποτε έλεγχο αισθητήρα  $\phi \in D(\Omega)$ ,

$$\langle v, \phi \rangle = (-1)^{\alpha_1 + \alpha_2} \langle u, \partial_1^{\alpha_1} \partial_2^{\alpha_2} \phi \rangle.$$

Εφόσον ο έλεγχος διαστήματος  $D(\Omega)$  είναι κλειστός κάτω από τη διαφοροποίηση, η νέα κατανομή ορίζεται σωστά.



**Εικόνα 3.1.** Οι εικόνες ως κατανομές ή γενικοποιημένες λειτουργίες. Τότε οι λειτουργικοί ελέγχοι μοντελοποιούν διάφορους βιολογικούς ή ψηφιακούς αισθητήρες, όπως οι αμφιβληστροειδικοί φωτοδέκτες της ανθρώπινης όρασης ή οι συσκευές διάταξης φορτίου στις κάμερες CCD.



Τα πολύ θετικά νέα σχετικά με την ιδέα των κατανεμιστικών εικόνων έχει ως εξής (βλέπε Εικόνα 3.1).

1. Οι γενικοποιημένες λειτουργίες φαίνεται να ικανοποιούν την προσδοκία των υψηλά διαφοροποιημένων σε αρκετές πειθαρχίες. Προσέξτε ότι ιστορικά η ιδέα των γενικοποιημένων λειτουργιών αναπτύχθηκε επίσης από την ανεπάρκεια κανονικών λειτουργιών στη μαθηματική ανάλυση.

2. Ο ορισμός μιας κατανομής φαίνεται να αντηχεί καλά με τη φυσική ερμηνεία μιας εικόνας. Με άλλα λόγια μια εικόνα  $u$  δεν υπάρχει από μόνη της στον κόσμο· αντιθέτως παρατηρείται μόνο ή εξασφαλίζεται μέσω των απαντήσεων του στους αισθητήρες (π.χ. κάμερες CCD ή το σύστημα οπτικής όρασης).

Ενδιαφέροντα παραδείγματα κατανεμιστικών εικόνων περιλαμβάνουν

1. ένα λαμπρό σημείο ιδανικά συγκεντρωμένο στη ρίζα (υποθέτοντας  $(0, 0) \in \Omega$ ):

$u(x) = u(x_1, x_2) = \delta(x)$ , όπου το  $\delta$  αντικαθιστά τη λειτουργία δέλτα του Ντιράκ ή δειγματικό σημείο, έτσι ώστε

$$\langle u, \phi \rangle = \phi(0, 0) \quad \text{για οποιοδήποτε αισθητήρα} \quad \phi \in D(\Omega);$$

2. μια ιδανικά ομοιόμορφη βηματική άκρη:  $u(x_1, x_2) = H(x_1)$ , όπου το  $H(t)$  είναι η βηματική συνάρτηση 0-1 βηματική λειτουργία·

3. μια ομοιόμορφη φωτεινή γραμμή (ή μια δέσμη [107]):  $u(x_1, x_2) = \delta(x_1)$ , έτσι ώστε  $\phi \in D(\Omega)$ , για οποιοδήποτε αισθητήρα

$$\langle u, \phi \rangle = \int_R \phi(0, x_2) dx_2.$$

Μια μη ομοιόμορφη φωτεινή γραμμή θα μπορούσε να μοντελοποιηθεί

σαν  $u(x_1, x_2) = \delta(x_1)v(x_2)$ , όπου  $v(x_2)$  αντιπροσωπεύει μια μονοδιάστατη κατανομή έτσι ώστε

$$\langle u, \phi \rangle = \langle v(x_2), \phi(0, x_2) \rangle.$$

Προσέξτε ότι οι εικόνες στα παραδείγματα 1 και 3 δεν μπορούν να περιγραφούν από τις κλασσικές λειτουργίες.

Το μειονέκτημα είναι ότι οι κατανεμιστικές εικόνες είναι τόσο ευρέες που κάποιος μετά δυσκολίας θα μπορούσε να πάρει οποιαδήποτε εντυπωσιακά χαρακτηριστικά εικόνας από αυτά.

Μια ενδιαφέρουσα αξιοσημείωτη συνθήκη για την ανάλυση εικόνων είναι η *θετικότητα*.

**Ορισμός 3.1 (Θετική Κατανομή).** Μια κατανεμιστική εικόνα  $u$  λέγεται ότι είναι θετικό αν για οποιοδήποτε αισθητήρα  $\phi \in D(\Omega)$  με  $\phi(x) \geq 0$ , κάποιος έχει  $\langle u, \phi \rangle \geq 0$ .

Το αξιοσημείωτο αποτέλεσμα της θετικότητας είναι η ιδιοκτησία συνέχειας.

**Θεώρημα 3.2.** Ας υποθέσουμε ότι το  $u$  είναι μια θετική κατανεμιστική εικόνα στο  $\Omega$ . Τότε για οποιοδήποτε συμπαγές σύνολο  $K \subseteq \Omega$ , υπάρχει μια σταθερά  $C = C_K(u) > 0$ , έτσι ώστε για οποιοδήποτε αισθητήρα  $\phi \in D(\Omega)$  με,  $\text{supp} \phi \subseteq K$ ,

$$\langle u, \phi \rangle \leq C \|\phi\|_\infty.$$

Η απόδειξη είναι ασήμαντη, καθώς κάποιος πάντα μπορεί να δημιουργήσει ένα μη-αρνητικό

αισθητήρα  $\psi \in D(\Omega)$  ο οποίος ισούται με 1 στο  $K$ . Ως αποτέλεσμα,  $\|\phi\|_\infty \psi - \phi$  είναι επίσης ένας μη-αρνητικός αισθητήρας για οποιοδήποτε  $\phi$  στηριζόμενο στο  $K$ .

Η θετικότητα απαραίτητα σημαίνει ότι η κατανομή δεν πρέπει να περιλαμβάνει οποιαδήποτε παράγωγα ενός αισθητήρα, επειδή η μονάδα  $\|\phi\|_\infty$  δεν ελέγχει τα παράγωγά της. Ως αποτέλεσμα, για παράδειγμα, το  $v(x_1) = \delta'(x_1)$  δεν μπορεί να είναι μια θετική μονοδιάστατη κατανομή.

Ένα πολύ ενδιαφέρον αποτέλεσμα από τη θετικότητα είναι το διάσημο *θεώρημα αναπαράστασης του Ριέζ* (βλέπε, π.χ., 118, 126).

**Θεώρημα 3.3 (Το Θεώρημα Αναπαράστασης του Ριέζ και οι Μετρήσεις Ραδονίου).** Το  $u$  είναι μια θετική κατανεμιστική εικόνα μόνο και μόνο αν δεν υπάρχει μια (μοναδική) μέτρηση Ραδονίου  $\mu$

στο  $\Omega$ , τέτοια που για οποιοδήποτε αισθητήρα  $\phi \in D(\Omega)$ ,

$$\langle u, \phi \rangle = \int_{\Omega} \phi(x) d\mu.$$

Εφόσον το  $\Omega$  είναι ένα υποσύνολο του  $\mathbb{R}^2$ , μια μέτρηση Ραδονίου  $\mu$  είναι απλά μια μέτρηση Μπόρελ η οποία είναι πεπερασμένη σε οποιοδήποτε συμπαγές υποσύνολο [118, 126]. Συνεπώς, ένα ιδανικό φωτεινό σημείο και μια ομοιόμορφη δέσμη στα προηγούμενα παραδείγματα είναι εξίσου θετικές κατανεμιστικές εικόνες.

Φυσικά, αν κάποιος αντιμετωπίζει τις σημειακές αξίες μιας εικόνας ως τους αριθμούς φωτονίων που προβάλλονται στα μικροστοιχεία, η συνθήκη θετικότητας γίνεται τότε πολύ φυσική.

Τώρα συζητούμε πώς να ποσοτικοποιήσουμε το σωστά το περιεχόμενο πληροφοριών μιας κατανομιστικής εικόνας. Έστω ότι το  $U$  υποδηλώνει οποιοδήποτε ανοιχτό υποσύνολο του  $\Omega$ . Ας ορίσουμε τη *συνολική μάζα* του  $u$  στο  $U$  από

$$\|u\|_U = \sup\{\langle u, \phi \rangle : \phi \in D(\Omega); \text{supp}\phi \subseteq U; \|\phi\|_\infty \leq 1\}.$$

Γενικά, αυτό θα μπορούσε να είναι  $\infty$ . Πρέπει να είναι πεπερασμένο όταν το  $U$  είναι προσυμπαγές (δηλαδή η σχετική του κλειστότητα στο  $\Omega$  είναι συμπαγές) και το  $u$  είναι θετικό.

Προσέξτε ότι η μέτρηση της συνολικής μάζας δεν μεταφέρει τη *διαφοροποιημένη* φύση. Δηλαδή, μια σταθερή εικόνα όπως  $u \equiv 1$  στο  $\Omega$  έχει ένα περιεχόμενο μη-μηδενικής πληροφορίας:

$\|u\|_\Omega = |\Omega|$ . Αυτό είναι ακατάλληλο από άποψης οράσεως επειδή πολλοί οπτικοί νευρώνες στον αρχικό φλοιό είναι γνωστό ότι ανταποκρίνονται μόνο στις χωρικές αλλαγές και χαρακτηριστικά (βλέπε, π.χ., τη δουλειά των βραβευμένων με Νόμπελ Χιουμπλ και Βίζελ [152]). Για βελτίωση, μπορούμε να ορίσουμε την ακόλουθη μέτρηση πληροφορίας, περιορίζοντας τους αισθητήρες σε εκείνους που είναι διαφοροποιημένοι:

$$\|u\|_U^* = \sup\{\langle u, \phi \rangle : \phi \in D(\Omega); \text{supp}\phi \subseteq U; \|\phi\|_\infty \leq 1; \langle 1, \phi \rangle = 0\}.$$

Η ολική ακεραιότητα του μηδέν στο  $\Omega$  σημαίνει απαραίτητα ότι το  $\phi$  πρέπει να είναι ταλαντωμένο. Έτσι υπό κάποια έννοια αυτά είναι αισθητήριες κυματιδίων.

Είναι τότε ξεκάθαρο ότι για οποιαδήποτε σταθερή εικόνα χωρίς χαρακτηριστικά  $u$ , το περιεχόμενο πληροφορίας  $\|u\|_U^* = 0$ . Από την άλλη, έχει παραμείνει ακόμη ένας βαθμός ελευθερίας έστω και στην νέο ορισμό, δηλαδή η επιλογή της τάξης διαφοροποιημένων αισθητήρων. Αν υποκατασταθεί από την τάξη

$$\{\phi \in D(\Omega) : \phi = \nabla \cdot \mathbf{g}, \mathbf{g} = (g_1, g_2) \in D(\Omega) \times D(\Omega), \|\mathbf{g}\|_\infty \leq 1\},$$

τότε καταλήγουμε με μια μονάδα  $\|\nabla u\|_U^*$  (ο τελεστής κλίσης είναι με την κατανομιστική έννοια), η οποία είναι ακριβώς η τηλεοπτική μέτρηση  $|Du|(U)$  if  $u$  is in  $L^1_{loc}(\Omega)$  (βλέπε ενότητα 2.2).

### 3.2.2 Εικόνες $L^p$

Οι εικόνες  $L^p$  μεταφέρουν περισσότερες δομές από τις κατανομιστικές ή τις εικόνες μέτρησης

Ραδονίου. Ας ανακαλέσουμε ότι για οποιοδήποτε  $p \in [0, \infty]$ , η λειτουργία διαστήματος

$L^p$  του Λεπέσκ ορίζεται ως

$$L^p(\Omega) = \left\{ u : \int_{\Omega} |u(x)|^p dx < \infty \right\}.$$

Όταν  $p = \infty$ , μια εικόνα  $L^\infty$  κατανοείται ως μια ουσιαστικά περιορισμένη λειτουργία. Αυτά είναι όλα διαστήματα του Μπάναχ κάτω από τις μονάδες

$$\|u\|_p = \left[ \int_{\Omega} |u|^p dx \right]^{1/p}.$$

Η σημαντική ιδιότητα των εικόνων  $L^p$  εκφράζονται για άλλη μια φορά από το θεώρημα

αναπαράσταση του Ριέζ. Ας θυμηθούμε ότι το  $p$  και το  $p^*$  λέγεται ότι κλίνουν αν

$$\frac{1}{p} + \frac{1}{p^*} = 1.$$

**Θεώρημα 3.4 (Το Θεώρημα Αναπαράστασης του Ριέζ σε Εικόνες  $L^p$ ).** Ας υποθέσουμε ότι

$1 \leq p < \infty$ . Τότε το δυαδικό του  $L^p(\Omega)$  είναι  $L^{p^*}(\Omega)$ . Με άλλα λόγια,

οποιαδήποτε συνεχόμενη γραμμική λειτουργικότητα  $L$  στο  $L^p(\Omega)$  πρέπει να δίνεται από τον τύπο ολοκληρώματος

$$L(u) = \int_{\Omega} v_L(x)u(x)dx$$

για κάποιο μοναδικό  $v_L \in L^{p^*}(\Omega)$ . Επιπλέον, η λειτουργική μονάδα του

$L$  και η μονάδα  $L^{p^*}$  του αναπαραστατικού  $v_L$  είναι πανομοιότυπα:

$$\|L\| = \|v_L\|_{p^*}.$$

Για  $p = \infty$ , παρόλο που το  $L^1$  είναι ισομετρικά ενσωματωμένο στο δυαδικό διάστημα  $(L^\infty)^*$ , το  $(L^\infty)^*$  είναι παρόλα αυτά πολύ μεγαλύτερο [126, 193].

Εφαρμόζοντας την ανισότητα του Χόλντερ,

$$\left| \int_{\Omega} f(x)g(x)dx \right| \leq \|f\|_q \|g\|_{q^*},$$

στο  $f = |u|^p, g = 1$ , και  $q = p_+/p$ , κάποιος καταλήγει ότι για

οποιοδήποτε  $p_+ \geq p \geq 1$ ,

$$L^{p_+}(\Omega) \subseteq L^p(\Omega),$$

εφόσον ο τομέας  $\Omega$  της εικόνας έχει υποτεθεί ότι είναι περιορισμένη. Με αυτή την έννοια, οι εικόνες  $L^1$  είναι η πιο γενική τάξη ανάμεσα σε όλα τα  $L^p$ 's.

Οι εικόνες  $L^p$  είναι φυσικά κατανομιστικές εικόνες. Εφόσον το  $D(\Omega)$  είναι πυκνό στο  $L^{p^*}$  για το  $1 \leq p^* < \infty$ , οποιαδήποτε εικόνα  $L^p$  πρέπει να είναι εντελώς καθορισμένη από την κατανομιστική συμπεριφορά στο  $D'(\Omega)$ . Αλλά οι εικόνες  $L^p$  μεταφέρουν περισσότερες δομές από τις γενικές κατανομιστικές εικόνες.

Μια ιδιαίτερα ενδιαφέρουσα δομή των εικόνων  $L^p$  είναι η εκρηκτική τους συμπεριφορά.

Εφαρμόζοντας το θεώρημα του Φουμπίνι στο  $g(t, x) = pt^{p-1} 1_{|u(x)| > t}$  με τη μέτρηση

προϊόντος  $dx \times dt$  στο  $\Omega \times (0, \infty)$ , κάποιος θα μπορούσε να εξασφαλίσει άμεσα την ακόλουθη αξιοσημείωτη ταυτότητα:

$$\int_{\Omega} |u(x)|^p dx = p \int_0^\infty t^{p-1} |\{x : |u(x)| > t\}| dt,$$

όπου η εγγραφή  $|A|$  στα δεξιά υποδηλώνει τη μέτρηση ενός συνόλου A του Λεβέσκ. Αυτό αποκαλείται συχνά εύθυμα ως η «αναπαράσταση στρώματος κέικ» (βλέπε, π.χ., Λιπ και Λος [193]).

Εφόσον το  $\Omega$  είναι περιορισμένο, το  $w(t) = |\{x : |u(x)| > t\}|$  είναι επίσης περιορισμένο. Έτσι η σύγκλιση στη δεξιά πλευρά της αναπαράστασης στρώματος κέικ (3.3) καθορίζεται από το ρυθμό μείωσης του  $w(t)$  ως  $t \rightarrow \infty$ .

1. Ως υποπροϊόν, και πάλι λαμβάνουμε  $L^{p+}(\Omega) \subseteq L^p(\Omega)$  για κάθε  $p+ \geq p \geq 1$ .

2. Ας υποθέσουμε ότι μια συγκεκριμένη εικόνα  $L^p u$  έχει ένα δυναμικό νόμο μείωσης  $w(t) = |\{x : |u(x)| > t\}| = O(t^{-\alpha})$ . Τότε πρέπει να έχουμε  $\alpha > p$ .

Τώρα συζητούμε πως να μετρήσουμε κατάλληλα το περιεχόμενο πληροφορίας μιας γενικής εικόνας  $L^p$  Δεδομένου οποιουδήποτε υποσυνόλου  $U \subset \Omega$  με θετική μέτρηση Λεπένσκ  $|U|$ , ας ορίσουμε το μέσο  $\langle u \rangle_U$  του  $u$  από

$$\langle u \rangle_U = \frac{1}{|U|} \int_U u(x) dx.$$

Το (μέσο) περιεχόμενο πληροφορίας του  $u$  στο  $U$  ορίζεται τότε ως το μέσο  $p$ -ταλάντωσης  $\sigma_p(u|U)$ :

$$\sigma_p(u|U) = \left( \frac{1}{|U|} \int_U |u - \langle u \rangle_U|^p dx \right)^{1/p}. \quad (3.4)$$

Όταν

$p = 2$ , αυτός είναι ο κανονικός ορισμός για εμπειρική σταθερή απόκλιση στη στατιστική. Επιπλέον, εφαρμόζοντας την ανισότητα του Χόλντερ (3.2), κάποιος μπορεί εύκολα να αποδείξει ότι

$$\sigma_p(u|U) \leq \sigma_q(u|U) \quad \text{for } 1 \leq p \leq q.$$

Σαν μια ενδιαφέρουσα σημείωση στην πραγματική ανάλυση, οποιαδήποτε εικόνα  $u$  με

$$\sup_{\text{any cube } U \subseteq \Omega} \sigma_p(u|U) < \infty$$

λέγεται ότι έχει περιορισμένες ταλαντώσεις μέσου- $p$ , και όλες οι εικόνες υποδηλώνονται από το

$BMO(\Omega)$ . Καταλύουμε ότι το διάστημα του  $BMO$  είναι ανεξάρτητο από το  $p \geq 1$  [216, 324].

### 3.2.3 Οι Εικόνες $H^n(\Omega)$ του Σοπόλεβ

Μια εικόνα  $L^2 u$  στο  $\Omega$  είναι φυσικά μια κατανέμηση όταν το εσωτερικό προϊόν του Χίλμπερτ εφαρμόζεται στο  $u$  και στους αισθητήρες ελέγχου στο  $D(\Omega)$ . Συνεπώς τα κατανεμιστικά παράγωγα  $\partial_1^{\alpha_1} \partial_2^{\alpha_2} u$  είναι καλά προσδιορισμένα.

Μια τέτοια εικόνα  $u$  πιστεύεται ότι είναι Σοπόλεβ  $H^1$  αν τα παράγωγα (σαν κατανομές) της πρώτης τάξης της ανήκουν επίσης στο  $L^2(\Omega)$  ή ισοδύναμα σε συνδυασμό, η κλίση του  $\nabla u$  ανήκει στο  $L^2 \times L^2$ . Με άλλα λόγια υπάρχουν δυο λειτουργίες  $L^2, u_1$  και  $u_2$ , έτσι ώστε για οποιοδήποτε έλεγχο φορέα αισθητήρα  $\phi = (\phi_1, \phi_2) \in D(\Omega) \times D(\Omega)$ ,

$$\langle \nabla u, \phi \rangle = \langle \partial_1 u, \phi_1 \rangle + \langle \partial_2 u, \phi_2 \rangle \equiv - \int_{\Omega} (u_1 \partial_1 \phi_1 + u_2 \partial_2 \phi_2) dx. \quad (3.5)$$

Το

$H^1(\Omega)$  είναι ένα διάστημα Χίλμπερτ κάτω από το εσωτερικό προϊόν,

$$\langle u, v \rangle_{H^1} = \langle u, v \rangle_{L^2} + \langle \nabla u, \nabla v \rangle_{L^2 \times L^2},$$

το οποίο προκαλεί τη γραμμική μονάδα

$$\|u\|_{H^1} = \sqrt{\|u\|_{L^2}^2 + \|\nabla u\|_{L^2 \times L^2}^2}.$$

Για τις εικόνες  $H^1$ , φαίνεται φυσικό να μετρήσουμε τις πληροφορίες εικόνας από το

$$\|\nabla u\|_{L^2 \times L^2} = \left( \int_{\Omega} ((\partial_1 u)^2 + (\partial_2 u)^2) dx \right)^{1/2}.$$

Τα διαστήματα εικόνων Σοπόλεβ υψηλότερης τάξης  $H^n(\Omega)$  ( $n = 2, 3, \dots$ ) προσδιορίζονται με παρόμοιο τρόπο και μεταφέρουν τα φυσικά διαστήματα δομών του Χίλμπερτ

όπως στο (3.5). Από την άλλη, αν η μονάδα  $L^2$  που χρησιμοποιείται για να περιορίσει τα παράγωγα αντικατασταθεί από μια μονάδα  $L^p$  ( $p \geq 1$ ), τα συνιστάμενα διαστήματα εικόνων του Σοπόλεβ υποδηλώνονται από  $W^{n,p}(\Omega)$  [3, 126, 193].

### 3.2.4 Εικόνες BV

Από γενικές κατανομής στις εικόνες  $L^p$ , και περαιτέρω στις εικόνες Σοπόλεβ, κάθε φορά όταν προστίθενται περισσότερες κανονικότητες, τα διαστήματα γίνονται πιο συγκεκριμένα και πιο ευχέριστα. Παρόλα αυτά υπάρχει ένα δίλημμα ανάμεσα στην κανονικότητα και την πιστότητα της μοντελοποίησης εικόνων. Μια ιδανική λύση είναι να βρούμε την ισορροπία, δηλαδή, να δημιουργήσουμε ένα διάστημα εικόνας που θα είναι και μαθηματικά ελατό και και πρακτικά πιστά στην αναπαράσταση χαρακτηριστικών κλειδί των γενικών εικόνων. Αυτό είναι το διάστημα των εικόνων BV όπως αναφέρθηκε ήδη στο προηγούμενο κεφάλαιο, του οποίου ο ρόλος στην ανάλυση και επεξεργασία εικόνων υπήρξε ουσιαστική και εμπνευστική από τη πασίγνωστη δουλειά των Ράτιν Όσερ [257] και των Ρατιν, Όσερ και Φατέμι [258].

Από την άποψη της κατανομής, μια εικόνα  $L^1 u$  ανήκει στο  $BV(\Omega)$  μόνο και μόνο αν η κατανομιστική κλίση  $\nabla u$  ικανοποιεί

$$\sup_{\phi \in D(\Omega) \times D(\Omega): \|\phi\|_\infty \leq 1} \langle \nabla u, \phi \rangle < \infty,$$

όπου  $\|\phi\|_\infty = \sup_{x \in \Omega} (\phi_1^2(x) + \phi_2^2(x))^{1/2}$ . Αν ισχύει αυτό, αυτός ο υπολογισμός

ονομάζεται η τηλεόραση του  $u$  και υποδηλωμένο από το  $TV[u]$  ή  $|Du|(\Omega)$ . Το  $BV(\Omega)$  είναι ένα διάστημα του Μπάναχ κάτω από τη φυσική μονάδα

$$\|u\|_{BV} = \|u\|_{L^1} + TV[u].$$

Τα αξιοσημείωτα χαρακτηριστικά των εικόνων BV διακρίνονται ως ακολούθως.

1. όλες οι εικόνες  $W^{1,1}$  (συμπεριλαμβανομένης  $H^1$  ως υποδιαστήματος εφόσον το  $\Omega$  είναι περιορισμένο) είναι εικόνες BV, αλλά οι πρώτες δεν επιτρέπουν άκρα (δηλαδή άλματα έντασης κατά μήκος των καμπύλων). Από την αρχή της έρευνας στην τεχνητή νοημοσύνη, τα άκρα αναγνωρίστηκαν ως βασικά οπτικά σημεία στην αντίληψη και κατανόηση των εικόνων.

2. Εφόσον οι εικόνες  $L^p$  επιτρέπουν επίσης τα άκρα, δεν σημειώνουν τοπικές ταλαντωτικές ανωμαλίες. Αλλά οι εικόνες BV το κάνουν.

Έτσι η ιδέα των εικόνων BV επιτυγχάνει πράγματι μια λογικά καλή ισορροπία ανάμεσα στην καταστολή ανωμαλιών (συχνά λόγω θορύβου) και το σεβασμό ενδογενών χαρακτηριστικών εικόνας όπως τα άκρα.



Η ιδέα των εικόνων BV είναι επίσης συμβατή με τη συρρικνωμένη-γενικοποιημένη επεξεργασία εικόνων των Λι, Μάμφορτ και Χουάγκ [191]. Τα περισσότερα αντικείμενα στον τρισδιάστατο κόσμο

δεν είναι διάφανα, προκαλώντας το γενικό φαινόμενο της συρρίκνωσης. Στον τομέα  $\Omega$  δυσδιάστατων εικόνων, η συρρίκνωση μοντελοποιείται ως ακολούθως. Ας υποθέσουμε ότι υπάρχουν

$N$  υποτομείς  $\Omega$  του Λιπσχιτζ, ταξινομημένοι από το προσκήνιο στο παρασκήνιο:

$$Q_1, Q_2, \dots, Q_N,$$

των οποίων όλα τα όρια έχουν πεπερασμένες μονοδιάστατες μετρήσεις Χάουστοντοφ. Ας υποθέσουμε ότι το κάθε  $Q_n$  έχει ένα σταθερό απρόσωπο επίπεδο του  $u_n, n = 1 : N$ . Τότε η συνθετική εικόνα  $u$  δημιουργημένη από μια τέτοια σκηνή είναι

$$u(x) = u_1 1_{Q_1}(x) + u_2 1_{Q_2 \setminus Q_1}(x) + u_3 1_{Q_3 \setminus Q_2 \cup Q_1}(x) + \dots, \quad \text{δεδομέ}$$

νου ότι η ένταση του παρασκήνιου είναι 0. Τότε η τηλεόραση είναι

$$|Du|(\Omega) \leq \left( \max_{0 \leq i < j \leq N} |u_i - u_j| \right) \sum_{n=1}^N \mathcal{H}^1(\partial Q_n) < \infty, \quad (3.6)$$

υπονο

ώντας ότι μια τέτοια συνθετική εικόνα πρέπει να ανήκει στο  $BV(\Omega)$ . Αν τα  $u_n$  δικαιούνται να είναι ομαλές λειτουργίες αντί για σταθερές, η εικόνα αποτέλεσμα  $u$  είναι σημειακά ομαλή και ανήκει ακόμη στο  $BV(\Omega)$ .

### 3.3 Κυματίδια και Αναπαράσταση Πολλαπλής Κλίμακας

Σε αυτή την ενότητα, ακολουθώντας τη βασική γνώση που δόθηκε στο προηγούμενο κεφάλαιο [86, 96, 290], θα μελετήσουμε περαιτέρω το ρόλο των κυματιδίων σαν επαρκή εργαλεία για μοντελοποίηση και αναπαράσταση εικόνων πολλαπλής κλίμακας.

#### 3.3.1 Δόμηση Δυσδιάστατων Κυματιδίων

Πρώτα εισάγουμε δυο διαφορετικές προσεγγίσεις για τη δόμηση δυσδιάστατων κυματιδίων από μονοδιάστατα.

#### Προσέγγιση I: Τρία Μητρικά Κυματίδια σε Ίσες Κλίμακες

Όπως και στο προηγούμενο κεφάλαιο έστω  $\psi(t)$  και  $\phi(t)$  υποδηλώνουν το μητρικό κυματίδιο και τη λειτουργία κλιμάκωσης για μια ορθογωνική ανάλυση πολλαπλών διαστάσεων στο  $\mathbb{R}^1$ :

$$\dots \subseteq V_{-1} \subseteq V_0 \subseteq V_1 \subseteq \dots$$

έτσι ώστε το  $(\phi(t - k) \mid k \in \mathbb{Z})$  να είναι μια ορθογωνική βάση για το  $V_0$ , και το  $(\psi(t - k) \mid k \in \mathbb{Z})$  για το διάστημα κυματιδίων  $W_0$  το οποίο χαρακτηρίζει ορθογωνικά το  $V_0$  στο πιο «λεπτομερές» διάστημα  $V_1$ .

Τότε το

$$\{\psi_{i,k} = 2^{j/2} \psi(2^j t - k) \mid j, k \in \mathbb{Z}\}$$

είναι μια ορθογωνική βάση κυματιδίου για το  $L^2(\mathbb{R})$ .

Για να αποφευχθεί η πολυπλοκότητα, θα υποθεθεί σε αυτή την ενότητα ότι ο τομέας εικόνας

$\Omega = \mathbb{R}^2$ . Στην πράξη, υπάρχουν πολλοί τρόποι για να επεκταθούν τα δεδομένα εικόνας σε

έναν περιορισμένο τετράγωνο τομέα στο  $\mathbb{R}^2$ , όπως η μηδενική πλήρωση, η συμμετρική αντικατόπτριση, η περιοδική επέκταση, η προέκταση και άλλα (βλέπε π.χ., τους Στραγκ και Γκούγιεν [290]). Υπάρχουν επίσης αυστηροί τρόποι δόμησης κυματιδίων σε περιορισμένους τομείς (βλέπε π.χ., την πασίγνωστη δουλειά του Κόεν και άλλοι [89]).

Ας ανακαλέσουμε ότι ο το προϊόν τανυστής των δυο γραμμικών διαστημάτων  $X$  και  $Y$  καθορίζεται από το

$$X \otimes Y = \text{span}\{x \otimes y : x \in X \text{ and } y \in Y\}.$$

Αν και το  $X$  και το  $Y$  είναι διαστήματα Μπάναχ με άπειρες διαστάσεις, ο πιο πάνω ορισμός θα

έπρεπε να ακολουθείται από μια περαιτέρω διαδικασία ολοκλήρωσης έτσι ώστε  $X \otimes Y$  να είναι επίσης Μπάναχ.

Αν τόσο το  $X$  όσο και το  $Y$  είναι διαστήματα Χίλμπερτ, και  $(x_n \mid n \in \mathbb{Z})$  και  $(y_m \mid m \in \mathbb{Z})$  είναι ξεχωριστά οι ορθογωνικές τους βάσεις, τότε υπό τη φυσική επέκταση του προϊόντος τανυστή των εσωτερικών προϊόντων,

$$\langle x \otimes y, u \otimes v \rangle := \langle x, u \rangle \times \langle y, v \rangle,$$

τα προϊόντα τανυστές των βασικών φορέων

$$z_{n,m} = x_n \otimes y_m, \quad n, m \in \mathbb{Z},$$

γίνονται μια ορθογωνική βάση για  $X \otimes Y$ .

Η δόμηση του δυσδιάστατου MRA αρχίζει από την ακόλουθη σχέση.

**Θεώρημα 3.5 (Διαχωριστικότητα).** Αντιμετωπίζοντας  $\mathbb{R}^2 = \mathbb{R} \oplus \mathbb{R}$  (άμεσο άθροισμα διαστημάτων φορέων) κάποιος έχει υπό τις μετρήσεις Λεπέσκ,

$$L^2(\mathbb{R} \oplus \mathbb{R}) = L^2(\mathbb{R}) \otimes L^2(\mathbb{R}).$$

Ισχυρίζεται ότι για κάθε  $f(x) = f(x_1, x_2) \in L^2(\mathbb{R}^2)$ , και  $\varepsilon > 0$  υπάρχει μια λειτουργία  $f_\varepsilon(x)$  υπό τη μορφή του

$$f_\varepsilon(x) = \sum_{k=1}^K c_k g_k(x_1) e_k(x_2), \quad g_k, e_k \in L^2(\mathbb{R}), \quad \text{and } c_k \in \mathbb{R}$$

έτσι ώστε  $\|f - f_\varepsilon\|_{L^2(\mathbb{R}^2)} < \varepsilon$ . Υπάρχουν πολλές διαφορετικές προσεγγίσεις που οδηγούν σε αυτό το γεγονός και κανένα ίσως δεμ θα μπορούσε να είναι πιο βολικό για τον μετασχηματισμό του Φουριέ.

**Απόδειξη.** Προσέξτε ότι οι ελέγχοι λειτουργίας  $D(\mathbb{R}^2)$  (δηλαδή, ομαλές λειτουργίες με συμπαγές στήριξης) είναι πυκνές στο  $L^2(\mathbb{R}^2)$ . Για οποιαδήποτε λειτουργία  $\phi(x_1, x_2)$ , που υποθέτουμε ότι στηρίζεται στο κανονικό ανοιχτό τετράγωνο  $Q = (-\pi, \pi)^2$  χωρίς απώλεια γενικότητας, δημιουργούμε την περιοδική του επέκταση  $Q\Phi(x)$  έτσι ώστε  $\Phi(x)|_Q = \phi(x)$ . Προσέξτε ότι  $\Phi(x)$  ισούται με μηδέν σε ένα γειτονικό όριο του  $Q'$ . Τότε το  $\Phi(x)$  μπορεί να υπολογιστεί σε οποιαδήποτε ακρίβεια (στο  $L^2$  ή ομοιόμορφα) από οποιαδήποτε πεπερασμένη περικοπή της σειράς του Φουριέ

$$\sum_{(n,m) \in \Lambda} c_{n,m} e^{inx_1} e^{imx_2},$$

όπου το  $\Lambda \subseteq \mathbb{Z}^2$  υποδηλώνει κάποια κατάλληλη πεπερασμένη υποακολουθία. Ας ορίσουμε

$$g_n(x_1) = e^{inx_1} 1_{(-\pi, \pi)}(x_1) \in L^2(\mathbb{R}) \quad \text{and} \quad e_m(x_2) = e^{imx_2} 1_{(-\pi, \pi)}(x_2) \in L^2(\mathbb{R}).$$

Τότε

εύκολα συνεπάγεται ότι το  $\phi(x_1, x_2)$  μπορεί να υπολογιστεί σε οποιαδήποτε ακρίβεια στο

$$L^2(\mathbb{R}) \quad \text{από το βήμα του } g_n \otimes e_m \text{'s.} \quad \square$$

Για κάθε (λογαριθμική) κλίμακα  $j \in \mathbb{Z}$  ορίζουμε ένα κλειστό υποσύνολο του  $L^2(\mathbb{R}^2)$  από

$$V_{(j)} = V_j \otimes V_j = \text{closure}(\text{span}\{f_1(x_1)f_2(x_2) \mid f_1, f_2 \in V_j\}).$$

Εφόσον  $V_{j-1} \subseteq V_j$ , κάποιος πρέπει να έχει επίσης

$$V_{(j-1)} \subseteq V_{(j)}. \quad (3.9)$$

Η ομοιότητα διαστολής στο μονοδιάστατο MRA ισχύει επίσης σε δυο διαστάσεις:

$$f(x) \in V_{(j-1)} \Leftrightarrow f(2x) = f(2x_1, 2x_2) \in V_{(j)}. \quad (3.10)$$

Προσέξτε ότι οι παράγοντες διαστολής (δηλαδή, τα 2) είναι οι ίδιοι τόσο για το  $x_1$  όσο και για το  $x_2$ , το οποίο λέγεται ότι είναι ισοτροπικό. Τέλος, από το Θεώρημα 3.5 και την ολοκληρότητα του μονοδιάστατου MRA,

Κλειστότητα

$$\text{closure} \left( \lim_{j \rightarrow \infty} V_{(j)} \right) = L^2(\mathbb{R}^2). \quad (3.11)$$

Για να μελετήσουμε περαιτέρω την αναπαράσταση αυτών των διαστημάτων πολλαπλής ευκρίνειας, ορίζουμε

$$\begin{aligned} \psi^{(0,0)}(x) &= \phi(x_1) \otimes \phi(x_2), & \psi^{(0,1)}(x) &= \phi(x_1) \otimes \psi(x_2), \\ \psi^{(1,0)}(x) &= \psi(x_1) \otimes \phi(x_2), & \psi^{(1,1)}(x) &= \psi(x_1) \otimes \psi(x_2). \end{aligned} \quad (3.12)$$

Τότε το

$\psi^{(0,0)}$  είναι βαθυπερατό (δηλαδή, με συνολικό ολοκλήρωμα ίσο του 1) και είναι η δυσδιάστατη λειτουργία κλιμάκωσης, και τα υπόλοιπα τρία είναι όλα υψιπερατά (δηλαδή με μηδεν συνολικά ολοκληρώματα) και είναι τα δυσδιάστατα μητρικά κυματίδια.

Για οποιοδήποτε επίπεδο κλίμακας  $j \in \mathbb{Z}$ , και τοποθεσίας,  $k = (k_1, k_2) \in \mathbb{Z}^2$  ορίζουμε

$$\psi_{j,k}^\alpha(x) = 2^j \psi^\alpha(2^j x - \mathbf{k}), \quad (3.14)$$

όπου

$$\alpha = (a, b)$$

υποδηλώνει τους τέσσερις τύπους. Τότε το ακόλουθο θεώρημα είναι ξεκάθαρο.

**Θεώρημα 3.6 (Αποσύνθεση Δυσδιάστατων Κυματιδίων).**

1. Σε οποιαδήποτε κλίμακα  $j, (\psi_{j,k}^{(0,0)} \mid \mathbf{k} \in \mathbb{Z}^2)$  είναι μια ορθογωνική βάση για το  $V_{(j)}$ .

2. Έστω ότι το  $W_{(j)}$  υποδηλώνει το ορθογωνικό συμπλήρωμα του  $V_{(j)}$  in  $V_{(j+1)}$ .

Τότε

$$\{\psi_{j,k}^\alpha : \mathbf{k} \in \mathbb{Z}^2, |\alpha| = a + b > 0\}$$

είναι μια ορθογωνική βάση για το  $W_{(j)}$ .

3. Έτσι το  $V_{(1)} = V_{(0)} \oplus W_{(0)}$  έχει δυο σύνολα ορθογωνικών βάσεων:

$(\psi_{1,k}^{(0,0)} \mid \mathbf{k} \in \mathbb{Z}^2)$  και το συνδυασμό των ορθογωνικών βάσεων από το  $V_{(0)}$  και το  $W_{(0)}$  που μόλις εδραιώθηκαν πιο πάνω. Συνδέονται από τις σχέσεις διπλής κλίμακας

$$\psi^\alpha = \sum_{\mathbf{k} \in \mathbb{Z}^2} h_k^\alpha \psi_{1,k}^{(0,0)}, \quad (3.15)$$

όπου το

$h^\alpha$ 's  $\alpha = (0, 0)$  είναι τέσσερα ψηφιακά φίλτρα, βαθυπερατού για  $\alpha = (0, 0)$  και υψιπερατού για τα άλλα τρία.

Υπολογιστικά, τα τέσσερα δυσδιάστατα φίλτρα  $h^\alpha$ 's είναι επίσης τα προϊόντα τελεστές των δυο μονοδιάστατων ψηφιακών φίλτρων που καθορίζουν τη λειτουργία κλιμάκωσης και το κυματίδιο στο μονοδιάστατο MRA. Συνεπώς, η συστοιχία φίλτρων που καθιστά εφικτό την πιο πάνω αλλαγή

βάσεων είναι επίσης προϊόν τανυστής. Πιο συγκεκριμένα, ας υποθέσουμε ότι το

$$u \in V_{(j+1)}$$

έχει την αναπαράσταση

$$u(x_1, x_2) = \sum_k c_{j+1,k}^{(0,0)} \psi_{j+1,k}^{(0,0)}(x_1, x_2).$$

Αφού περάσει από τη μονοδιάστατη *ανάλυση* η συστοιχία φίλτρων (δηλαδή, βαθυπερατά και υπιπερατά κανάλια με υποδειγματισμό: βλέπε Ενότητα 2.6) κατά μήκος της κατεύθυνσης  $k_1$  για

οποιοδήποτε σταθερό  $k_2, (c_{j+1,k}^{(0,0)} \mid k \in \mathbb{Z}^2)$

διαχωρίζεται σε δυο ενδιάμεσα πεδία συντελεστών:

$$c_{(j,j+1),k}^{(0,0)} \text{ and } c_{(j,j+1),k}^{(1,0)},$$

όπου η ομοιόμορφη κλίμακα δείκτη διαχωρίζεται σε  $(j, j+1)$  εφόσον η κατεύθυνση

$k_2$  εξακολουθεί να είναι στο επίπεδο κλίμακας του  $j+1$ . Έπειτα, αφού περάσει από την ίδια

μονοδιάστατη συστοιχία φίλτρων κατά μήκος της κατεύθυνσης  $k_2$  για κάθε σταθερό  $k_1$ , τα δυο

ενδιάμεσα πεδία διαχωρίζονται περαιτέρω στα ποθητά τέσσερα πεδία συντελεστών του  $u$  στις

βάσεις  $\{\psi_{j,k}^\alpha \mid k \in \mathbb{Z}^2, \alpha = (0|1, 0|1)\}$ :

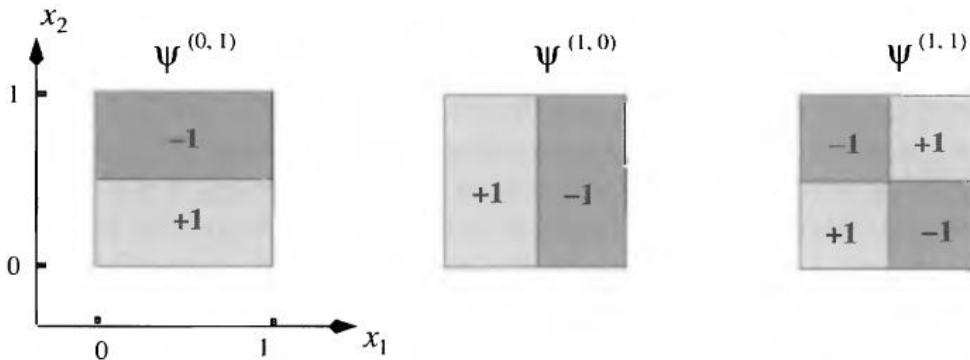
$$\{c_{j,k}^\alpha \mid k \in \mathbb{Z}^2\}, \quad \alpha = (0|1, 0|1).$$

**Παράδειγμα.** Σαν ένα παράδειγμα, θεωρήστε τα δυσδιάστατα κυματίδια του Χάαρ:

$$\psi^{(0,0)}(x) = 1_{[0,1]^2}(x_1, x_2),$$

$$\psi^{(0,1)}(x) = 1_{[0,1)}(x_1) \otimes (1_{[0,1/2)}(x_2) - 1_{[1/2,1)}(x_2)),$$

$$\begin{aligned} \psi^{(1,0)}(x) &= \psi^{(0,1)}(x_2, x_1), \\ \psi^{(1,1)}(x) &= (1_{[0,1/2)}(x_1) - 1_{[1/2,1)}(x_1)) \otimes (1_{[0,1/2)}(x_2) - 1_{[1/2,1)}(x_2)). \end{aligned} \quad (3.16)$$



**Εικόνα 3.2.** Τα τρία μητρικά κυματίδια του Χάαρ σε δυο διαστάσεις μέσω προϊόντων τανυστών.

Η εικόνα 3.2 βοηθά στην οπτικοποίηση αυτών των μητρικών κυματιδίων. Αυτό που φαίνεται πολύ ελκυστικό είναι ότι τα τρία κυματίδια μιμούνται τους τρεις διαφοροποιητικούς τελεστές (υπό το πεπερασμένο σχήμα διαφορετικότητας)

$$\partial_y, \quad \partial_x, \quad \partial_y \partial_x,$$

μέχρι κάποιους σταθερούς πολλαπλασιαστές.

Τέλος, ας προχωρήσουμε στον τομέα χωρικής συχνότητας για να κατανοήσουμε την ακριβή φύση τέτοιων δυσδιάστατων MRA.

Είτε στην περίπτωση των κυματιδίων Σάνον [96, 204, 216] ή από άποψης ασυμπτωματικής συμπεριφοράς της οικογένειας Ντόμπεχι των κυματιδίων [280, 281], το MRA σίγουρα παραλληλίζεται με τη θεωρία των Λίτλγουτ και Πάλεϋ για τη δυαδική τμηματοποίηση του τομέα του Φουριέ [127].

Έτσι στη μονοδιάστατη περίπτωση, κάθε διάστημα ευκρίνειας  $V_j$  μπορεί να θεωρηθεί περίπου ότι αναπαριστά όλα τα σήματα περιορισμένου εύρους στο  $|\omega| \leq 2^j \pi$  και το ορθογωνικό του

συμπλήρωμα  $W_j$  στο  $V_{j+1}$  όλα τα σήματα είναι περιορισμένου εύρους στη ζώνη ψηλής

συχνότητας  $|\omega| \in \pi(2^j, 2^{j+1}]$ . Κατά ακρίβεια, είναι η κύρια αλλαγή τις ιδέας της θεωρίας κυματιδίων να σκεφτούμε αυτές τις ζώνες ως ευκρίνεις κλίμακας αντί για ξεχωριστές συχνότητες [96, 204, 215].

Αυτή η άποψη επεκτείνεται φυσικά στο δυσδιάστατο MRA που μόλις δομήθηκε πιο πάνω. Το

μηχανισμός του προϊόντος τανυστή δηλώνει ότι κάθε ευκρίνεια  $V(j)$  αναπαριστά περίπου όλες τις δυσδιάστατες εικόνες περιορισμένου εύρους στο δυαδικό πρίσμα

$$Q_j = [-2^j \pi, 2^j \pi] \times [-2^j \pi, 2^j \pi],$$

και το συμπλήρωμα  $W(j)$  του κυματιδίου του υποδηλώνει τον Καρτεσιανό δαχτύλιο υψηλής

συχνότητας εντός του  $Q_{j+1}$  αλλά εκτός του  $Q_j$ .

Κάθε διάστημα  $W(j)$  κυματιδίου διαχωρίζεται περαιτέρω ορθογωνικά σε τρία υποδιαστήματα που αναπαριστούνται από τρία μητρικά κυματίδια

$$W(j) = W_{(j)}^{(0,1)} \oplus W_{(j)}^{(1,0)} \oplus W_{(j)}^{(1,1)},$$

που αναλογούν στην περαιτέρω τμηματοποίηση του Καρτεσιανού δαχτυλίου σε τρεις ζώνες:

$$\begin{aligned} R_j^{(0,1)} &= \{(\omega_1, \omega_2) : |\omega_1| \leq \pi 2^j, \pi 2^j < |\omega_2| \leq \pi 2^{j+1}\}, \\ R_j^{(1,0)} &= \{(\omega_1, \omega_2) : |\omega_2| \leq \pi 2^j, \pi 2^j < |\omega_1| \leq \pi 2^{j+1}\}, \\ R_j^{(1,1)} &= \{(\omega_1, \omega_2) : \pi 2^j < |\omega_1| \text{ and } |\omega_2| \leq \pi 2^{j+1}\}. \end{aligned} \quad (3.17)$$

### Προσέγγιση II: Μοναδικό Μητρικό Κυματίδιο και Μίξη Ανισοτροπικής Κλίμακας.

Ένα ιδιαίτερο χαρακτηριστικό της πιο πάνω δόμησης είναι ότι κάθε διάστημα  $V(j)$  αναλογεί σε ένα μοναδικό επίπεδο ευκρίνειας  $j$  τόσο στην κατεύθυνση  $x_1$  όσο και στην  $x_2$ . Υπάρχει, παρόλα αυτά, ένας άλλος τρόπος δόμησης βασισμένος στο προϊόν τανυστή που επιτρέπει τη μίξη ανισοτροπικής κλίμακας.

Επιπρόσθετα στη μετάφραση φορέα  $k = (k_1, k_2)$ , ορίζουμε

$j = (j_1, j_2) \in \mathbb{Z}^2$  ως τον φορέα ανισοτροπικής ευκρίνειας. Σύμφωνα με τη φόρμουλα ανισοτροπικής ευκρίνειας (3.7), τα προϊόντα τανυστές

$$\psi_{j,k}(x) = \psi_{j_1,k_1}(x_1) \otimes \psi_{j_2,k_2}(x_2), \quad j, k \in \mathbb{Z}^2,$$



συνίστανται σε μια ορθογωνική βάση για  $L^2(\mathbb{R}^2)$ . Για να ερευνήσουμε περαιτέρω τις ενδογενείς δομές αυτής της οικογένειας δυοδιάστατων κυματιδίων, για κάθε φορέα ανισοτροπικής ευκρίνειας,  $\mathbf{j} = (j_1, j_2)$ , ορίζουμε τον διαγώνιο πίνακα

$$J = \begin{pmatrix} j_1 & 0 \\ 0 & j_2 \end{pmatrix}.$$

Τότε από τη γραμμική άλγεβρα,

$$2^J = e^{J \ln 2} = \begin{pmatrix} e^{j_1 \ln 2} & 0 \\ 0 & e^{j_2 \ln 2} \end{pmatrix} = \begin{pmatrix} 2^{j_1} & 0 \\ 0 & 2^{j_2} \end{pmatrix}.$$

Αντιμετω

πίζοντας το μεταβλητό συνεχόμενο μικροστοιχείο,  $x \in \mathbb{R}^2$  σαν ένα φορέα στήλης

$$(x_1, x_2)^T \quad | \quad \mathbf{k} = (k_1, k_2)^T \quad \mathbf{j} \in \mathbb{Z}^2,$$

και επίσης, έχουμε τότε για κάθε

$$\psi_{\mathbf{j}, \mathbf{k}}(x) = \sqrt{\det 2^J} \psi_{\mathbf{0}, \mathbf{0}}(2^J x - \mathbf{k}),$$

το οποίο μοιάζει πιο πολύ με τη μονοδιάστατη φόρμουλα.

$$\mathbf{j} = (j_1, j_2),$$

Για κάθε φορέα ανισοτροπικής ευκρίνειας υποδιάστημα κυματιδίου ορίζουμε το συνδεδεμένο

$$W_j = \text{closure}(\text{span}\{\psi_{\mathbf{j}, \mathbf{k}} \mid \mathbf{k} \in \mathbb{Z}^2\}).$$

Ακολουθώντας τη συζήτηση στην προηγούμενη υποενότητα, στον τομέα του Φουριέ, το

$W_j$  αναλογεί στα τέσσερα ορθογώνια υψηλών συχνοτήτων:

$$Q_j = \{(\omega_1, \omega_2) : 2^{j_1} \pi < |\omega_1| \leq 2^{j_1+1} \pi \text{ and } 2^{j_2} \pi < |\omega_2| \leq 2^{j_2+1} \pi\}.$$

Αν

$|j_1 - j_2| \gg 1$ , το  $Q_j$  αναπαριστά ένα ορθογώνιο με μεγάλη σχέση αναλογίας, λύνοντας τα χαρακτηριστικά εικόνες που είναι σχετικά υπερβολικά ταλαντωτικά προς μια κατεύθυνση και πιο

ομαλά προς την άλλη. Έτσι κάποιος θα μπορούσε να φανταστεί ότι αυτή η δεύτερη δόμηση είναι ιδιαίτερα αποτελεσματική στην αναπαράσταση δομών εικόνων όπως ένα λεπτό μεταξένιο σύννεφο, μαλλιά ή γούνα, ψηλό γρασίδι, ή καλαμάκια. Στον τομέας φυσικής, τέτοιες δομές πρέπει να μεταφέρουν επίσης μεγάλες σχέσεις αναλογίας σύμφωνα με την αρχή αβεβαιότητας του Χάιζενμπεργκ [288].

### 3.3.2 Αντιδράσεις Κυματιδίων στα Τυπικά Χαρακτηριστικά Εικόνων

Σε αυτή την ενότητα συζητούμε τις αντιδράσεις κυματιδίων σε τυπικά χαρακτηριστικά εικόνων: ομαλότητα και άκρα. Για να δείξουμε τις κύριες ιδέες, θα περιορίσουμε την ανάλυση μόνο στη μονοδιάστατη περίπτωση.

#### Αντίδραση σε Ομαλές Εικόνες

Ένα μητρικό κυματίδιο  $\psi(x) \in L^2(\mathbb{R})$  θεωρείται  $r-1$  κανονικό αν ανήκει στο  $C^r(\mathbb{R})$ .

Για ορθογωνικό MRA, υπάρχει ένα όμορφο θεώρημα που συνδέει την ομαλότητα του μητρικού κυματιδίου με τον αριθμό στιγμών που εξαφανίζεται, ανήκει στους Μέγιερ [216], Μπατλ [18], και Ντόμπεχις [96]. Μια βελτίωση στο αποτέλεσμα του Ντόμπεχι έγινε επίσης από τους Κάι και Σιεν [38].

**Θεώρημα 3.7 (Στιγμες Εξαφάνισης [96, 38]).** *Ας υποθέσουμε ότι το μητρικό κυματίδιο  $\psi(x)$  είναι  $r-1$  κανονικό. Επιπλέον, το (i)  $\psi$  έχει μέχρι και στιγμέ τάξης και (ii)  $|\psi(x)| \leq C(1 + |x|)^r, x \in \mathbb{R}$ , για κάποια σταθερή σταθερά  $C$ . Τότε όλες οι στιγμές εξαφανίζονται μέχρι την τάξη  $r$ th*

$$\int_{\mathbb{R}} x^k \psi(x) dx = 0, \quad k = 0 : r. \quad (3.18)$$

Δείτε τους Ντόμπεχι [96] και τους Κάι και Σιεν [38] ως απόδειξη. Προσέξτε ότι η συνθήκες (i) και (ii) ικανοποιούνται μηδανικά για συμπαγώς στηριζόμενα κυματίδια, τα οποία είναι τα πιο συνηθισμένα στις εφαρμογές [95]. Συγκεκριμένα, έχουμε τα ακόλουθα αποτελέσματα.

#### Θεώρημα 3.8 (Μείωση Ρυθμού και Ομαλότητα Κυματιδίων).

(1) Ένα συμπαγώς στηριζόμενο μητρικό κυματίδιο  $\psi$  δεν μπορεί να είναι  $C^\infty$ .

(2) Πιο γενικά, ένα μητρικό κυματίδιο  $\psi$  που μειώνεται με συγκεκριμένο εκθετικό ρυθμό (δηλαδή,  $O(e^{-a|x|})$  για κάποιο  $a > 0$  at  $x = \infty$ ) (σε) δεν μπορεί να είναι  $C^\infty$ .

Το (1) είναι αποτέλεσμα του (2), ενώ η απόδειξη του (2) είναι εμφανής από το προηγούμενο θεώρημα: διαφορετικά όλες οι στιγμές του  $\psi$  θα εμφανίζονταν, ή ισοδύναμα, όλοι οι συντελεστές

Τείλορ του  $\hat{\psi}$  (μετασχηματισμός του Φουριέ) εξαφανίζεται στην αρχή, πράγμα αδύνατο καθώς το

$\psi$  είναι μια μη-μηδενική λειτουργία και αναλυτική σε μια οριζόντια ταινία που περιέχει ένα πραγματικό άξονα (λόγω της εκθετικής μείωσης).

Στη συνέχεια, θα υποθέσουμε ότι το μητρικό κυματίδιο  $\psi$  είναι  $r$ -κανονικό και συμπαγώς σφαιρικό σε ένα πεπερασμένο διάστημα  $[a, b]$ . Όσο για τα κυματίδια Ντόμπτσεϊ [96, 95], το  $r$  θα μπορούσε να είναι μια επιθυμητή ακέραιη τάξη.

Ας υποθέσουμε ότι το  $u(x)$  είναι οποιαδήποτε μονοδιάστατη εικόνα  $C^m$  με

$0 \leq m \leq r + 1$ . Αναμένουμε τους συντελεστές κυματιδίων να είναι μικροί σε μικρές κλίμακες (δηλαδή μεγάλα  $j$ ) όταν το σήμα συμπεριφέρεται ομαλά σε σχέση με τα μικροσκοπικά κυματίδια.

Κατά ορισμό, ο συντελεστής κυματιδίων  $c_{j,k}$  είναι

$$c_{j,k} = \int_{\mathbb{R}} u(x) \psi_{j,k}(x) dx = 2^{-j/2} \int_a^b \psi(y) u(x_0 + 2^{-j}y) dy, \quad (3.19)$$

όπου

$x_0 = x_{j,k}$  υποδηλώνει  $2^{-j}k$ . Κατά την διαστολή του Τείλορ στο  $x_0$ ,

$$u(x_0 + t) = \sum_{n=0}^{m-1} \frac{u^{(n)}(x_0)}{n!} t^n + \int_0^t \frac{(t-s)^{m-1}}{(m-1)!} u^{(m)}(x_0 + s) ds = P_{m-1}(t) + R_{m-1}(t).$$

Για

$$C_m = C_m(u) = \|u^{(m)}\|_\infty < \infty,$$

ευκολία ας υποθέσουμε ότι  $C_m < \infty$ . Τότε το κατάλοιπο

$R_m$

μπορεί να περιοριστεί εύκολα:

$$|R_{m-1}(t)| \leq \frac{C_m}{m!} |t|^m.$$

Από την άλλη, εφόσον  $r \geq m-1$ , το εσωτερικό προϊόν του  $\psi(y)$  και  $P_{m-1}(2^{-j}y)$  πρέπει να εξαφανιστεί.

Συνεπώς,

$$c_{j,k} = 2^{-j/2} \int_a^b \psi(y) R_{m-1}(2^{-j}y) dy.$$

Ορίζουμε την απόλυστη στιγμή της τάξης  $m$   $M_m$  του  $\psi$  από

$$M_m = M_m(\psi) = \int_a^b |\psi(y)| |y|^m dy.$$

Τότε έχουμε εδραιώσει στον υπολογισμό για τους συντελεστές κυματιδίων

$$|c_{j,k}| \leq 2^{-j(m+1/2)} \frac{C_m(u)}{m!} M_m(\psi) \text{ or } c_{j,k} = O(h_j^{m+1/2}), \quad (3.20)$$

όπου όπως και στις υπολογιστικές ΛΣΕ, το  $h_j$  υποδηλώνει την κλίμακα  $2^{-j}$  στο επίπεδο  $j$ .

Έστω ότι το  $D_j$  υποδηλώνει την ορθογώνια προβολή από το  $L^2$  στο διάστημα κυματιδίου

$$W_j, \text{ και } w_j(x) = D_j u(x).$$

Λόγω της περιορισμένης αλληλοκάλυψης μεταξύ των  $\psi_{j,k}$ 's σε κάθε σταθερή ευκρίνεια  $j$ , ο υπολογισμός στο (3.20) υπονοεί άμεσα ότι

$$\|w_j\|_\infty = O(h_j^m). \quad (3.21)$$

Προσέξ

τε ότι ο επιπλέον παράγοντας  $h_j^{1/2}$  που περιορίζει  $c_{j,k}$ 's στο (3.20) έχει ακυρωθεί από τον

παράγοντα κανονικοποίησης  $h_j^{-1/2}$  στο

$$\psi_{j,k}(x) = h_j^{-1/2} \psi(h_j^{-1}(x - x_{j,k})).$$

Και στις δυο φόρμουλες (3.20) και (3.21), το  $O(\cdot)$  μπορεί να βελτιωθεί σε  $o(\cdot)$  μετά από μικρή προσαρμογή του επιχειρήματος.

Εκτός από αυτή την κλασσική ή σημειακή προσέγγιση, υπολογισμοί όπως το (3.21) μπορούν να εδραιωθούν για τις εικόνες Σοπόμεβ σε ακόμη πιο βολικούς τρόπους, τους οποίους επεξηγούμε τώρα.

Από άποψης του μετασχηματισμού του Φουριέ  $\hat{u}(\omega)$  [193] η συνθήκη  $u \in H^m(\mathbb{R})$  είναι ίση με το

$$\int_{\mathbb{R}} |\hat{u}|^2 (1 + 4\pi^2 \omega^2)^m d\omega < \infty. \quad (3.22)$$

Ας

υποθέσουμε ότι  $\hat{u}$  έχει ένα δυναμικό νόμο μείωσης

$$\hat{u}(\omega) = O(|\omega|^{-\alpha})$$

για κάποιο ιδανικό  $\alpha$ . Τότε το (3.22) δηλώνει ότι (σε μια διάσταση)

$$\alpha - m > 1/2 \text{ or } \alpha > m + 1/2.$$

Ας υποθέσουμε ότι η κανονικότητα  $r$  του μητρικού κυματιδίου  $\psi$  είναι πολύ μεγαλύτερη από το  $m$ . Τότε στον τομέα του Φουριέ, το επίπεδο  $j$ th της συνιστώσας κυματιδίου  $w_j$  του  $u$  είναι

$$\alpha - m - 1/2 \quad \varepsilon > 0)$$

περίπου ίσο του (με δηλωμένο από

$$\hat{w}_j(\omega) \simeq \hat{u}(\omega) \cdot 1_{[2^j\pi, 2^{j+1}\pi]}(|\omega|) = O(|\omega|^{-(m+1/2+\varepsilon)}) \cdot 1_{[2^j\pi, 2^{j+1}\pi]}(|\omega|),$$

όπου στην πραγματικότητα ενδεικτική λειτουργία θα έπρεπε να αντικατασταθεί από μια ομαλή λειτουργία (άναλογα με τη δόμηση του συνδεδεμένου MRA [96, 216]). Συνεπώς, για κάποια σταθερά  $C$ .

$$\|w_j\|_{\infty} \leq C \|\hat{w}_j\|_{L^1} = O(h_j^{m-1/2+\varepsilon}). \quad (3.23)$$

Έτσι από την πλευρά της θεωρίας υπολογισμού, για ομαλές εικόνες, οι συντελεστές και οι συνιστώσες κυματιδίων υψηλής ευκρίνειας μπορούν όλα να τεθούν στο 0.

### Αντίδραση Στα Άκρα Εικόνες

Ας θεωρήσουμε ένα κανονικοποιημένο άκρο της μονοδιάστατης εικόνας Χέβισαϊτ

$$u(x) = H(x) = 1_{[x_e, \infty)}(x),$$

όπου  $x_e$  είναι μικροστοιχείο άκρου ή άλματος. Όπως και στην προηγούμενη ενότητα, γράφουμε

$$x_{j,k} = 2^{-j}k \quad \text{και} \quad h_j = 2^{-j}, \quad \psi = [a, b].$$

και υποθέτουμε ότι  $\psi = [a, b]$ . Τότε ο συντελεστής κυματιδίου είναι

$$c_{j,k} = \langle u, \psi_{j,k} \rangle = h_j^{1/2} \int_a^b \psi(y) u(x_{j,k} + h_j y) dy. \quad (3.24)$$

Αν το

μικροστοιχείο άκρου  $x_e \notin x_{j,k} + h_j [a, b]$ , κάποιος έχει  $c_{j,k} = 0$  λόγω της σταθερότητας της εικόνας Χέβισαϊτ μακριά από το άκρο της. Διαφορετικά, ορίζοντας

$$y_{j,k} = (x_e - x_{j,k}) / h_j,$$

κάποιος έχει

$$c_{j,k} = h_j^{1/2} \int_{y_{j,k}}^b \psi(y) dy \leq h_j^{1/2} \|\psi\|_{L^1}.$$

Εφόσον  $y_{j,k} - y_{j,k+1} = 1$ , υπάρχουν το πολύ  $N = [b - a] + 1$  αριθμοί μη

μηδενικών συντελεστών κυματιδίων σε κάθε επίπεδο ευκρίνειας  $j$ . Προσέξτε ότι το  $N_j$  είναι

ανεξάρτητο από το  $j$ , έτσι στο πλάνο διαστήματος – κλίμακας (δηλαδή  $x$  ενάντια  $\log_2 h$  σε

πλάνο με  $x = 2^{-j}k$  και  $h = 2^{-j}$ ), ο τομέας επίδρασης τέτοιου ιδανικού άκρου φέρει το σχήμα ομπρέλας [96, 290].

Πιο γενικά, αν το σήμα της μονοδιάστατης εικόνας  $u$  είναι σημειακά ομαλό και έχει πολλά σημεία άκρου, τότε κάθε σημείο άκρου μεταφέρει τον δικό του τομέα επίδρασης ομπρέλας, και αλληλοκαλύπτονται από πλατιές ευκρίνειες. Μέσα σε κάθε ομπρέλα, οι συντελεστές κυματιδίου είναι περιορισμένοι από

$$|c_{j,k}| \leq h_j^{1/2} \|u\|_\infty \|\psi\|_1, \quad (3.25)$$

ενώ όλοι οι υπόλοιποι παραμένουν υποκείμενοι στον υπολογισμό που αναπτύχθηκε στην προηγούμενη ενότητα.

### 3.3.3 Οι Εικόνες του Μπεσόβ και η Αναπαράσταση Ελλιπών Κυματιδίων

Σαν ένα εργαλείο πολλαπλής κλίμακας, τα κυματίδια είναι ιδιαίτερα δυναμικά για τη μελέτη μιας τάξης εικόνων γνωστών ως εικόνων Μπεσόβ, των οποίων η φύση πολλαπλής κλίμακας είναι ενδογενής από τους ορισμούς τους.

Μια τάξη Μπέσοβ στο  $\mathbb{R}$  είναι συχνά υποδηλούμενη από  $B_q^\alpha(L^p)$ , με τρεις δείκτες  $\alpha$ ,  $p$ , και

$q$ :

1. το  $\alpha$  αποκαλείται ο δείκτης (ή δύναμη) κανονικότητας, μετρώντας το βαθμό ομαλότητας.
2. το  $L^p = L^p(\mathbb{R})$  υποδηλώνει τη μετρική ενδοκλίμακα που ελέγχει τις πεπερασμένες αλλαγές σε κάθε κλίμακα. Θα πρέπει να αποκαλέσουμε τον δείκτη ενδοκλίμακας  $p$
3. Το  $q$ , ή  $L^q(dh/h, \mathbb{R}^+)$ , υποδηλώνει την μετρική διακλίμακα (στο διάστημα κλίμακας

$h \in \mathbb{R}^+$  με τη λογαριθμική μέτρηση  $dh/h$ ) που ελέγχει τη γενική κανονικότητα δια μέσου

όλων των κλιμάκων. Έτσι αποκαλούμε το  $q$  τον δείκτη διακλίμακας.

Ακολουθως, θα αρχίσουμε με τον κλασσικό ορισμό των διαστημάτων του Μπεσόβ και έπειτα θα δηλώσουμε το αντίστοιχο υπό τις αναπαραστάσεις κυματιδίων. Για απλότητα, επικεντρωνόμαστε στη μονοδιάστατη περίπτωση όταν ο τομέας εικόνας είναι  $\Omega = \mathbb{R}$ .

#### Εικόνες Μπέσοβ και Χαρακτηρισμός Πολλαπλής Κλίμακας

Θα υποθεθεί καθόλη την ενότητα ότι ο δείκτης κανονικότητας  $\alpha \in [0, 1)$  και

$u(x) \in L^p(\mathbb{R})$ . Για οποιαδήποτε κλίμακα  $h > 0$ , ορίζουμε τον συντελεστή  $p^{-1}$  της

συνέχειας του  $u$  να είναι

$$\omega_p(u, h) = \sup_{|a| \leq h} \|u(x+a) - u(x)\|_{L^p(\mathbb{R})}. \quad (3.26)$$

Είναι

ομοιόμορφα περιορισμένο από το  $2\|u\|_{L^p}$  λόγω τις τριγωνικής φόρμουλας των μονάδων

$L^p$  Μια εικόνα  $u$  λέγεται να ανήκει στην τάξη Μπέσοβ  $B_q^\alpha(L^p)$  αν

$$\left[ \int_0^\infty \frac{\omega_p(u, h)^q}{h^{\alpha q}} \frac{dh}{h} \right]^{\frac{1}{q}} < \infty.$$

Αν ισχύει αυτό, η ομοιογενής (ημι)μονάδα Μπέσοβ του  $B_q^\alpha(L^p)$  ορίζεται ως

$$\|u\| = \left[ \int_0^\infty \frac{\omega_p(u, h)^q}{h^{\alpha q}} \frac{dh}{h} \right]^{\frac{1}{q}} = \|\omega_p(u, h)h^{-\alpha}\|_{L^q(dh/h)}, \quad (3.27)$$

και η

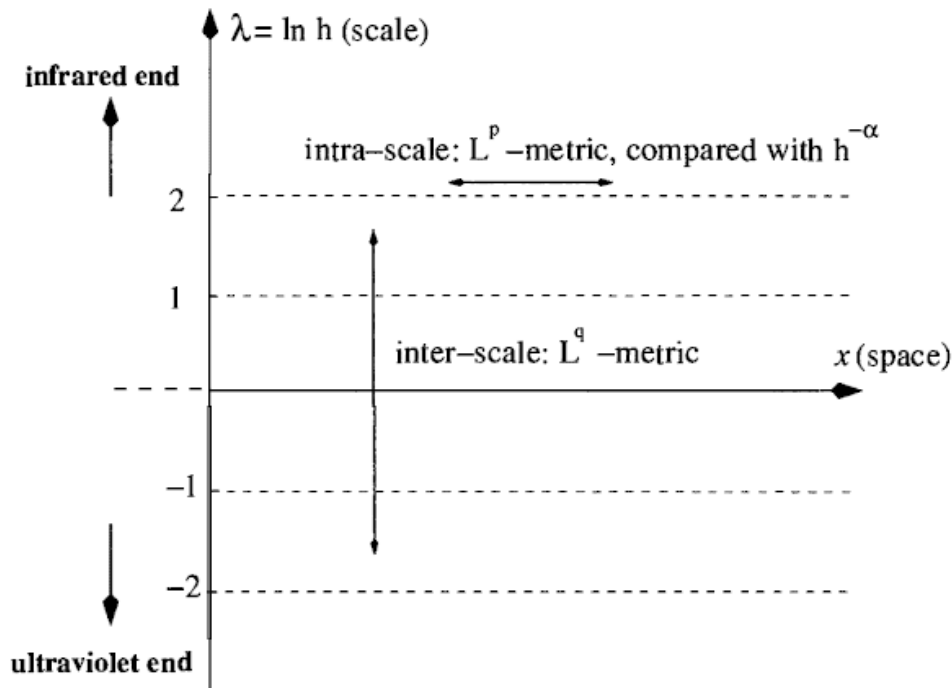
πραγματική μονάδα του  $B_q^\alpha(L^p)$  ορίζεται από

$$\|u\|_{B_q^\alpha(L^p)} = \|u\|_p + \|u\|. \quad (3.28)$$

Όπως και στην κλασική θεωρία  $L^p$  αν είτε το  $p < q$  ή το  $q$  είναι  $\infty$ , η αντίστοιχη μονάδα γίνεται κατανοητή με την έννοια της ουσιαστικής εξίσωσης [126, 193]. Η Εικόνα 3.3 βοηθά στην

οπτικοποίηση της σημασίας των τριών παραμέτρων  $\alpha, p, q$  και με τον ορισμό των μονάδων Μπέσοβ.





Εικόνα

3.3 Οι Μονάδες Μπέσοβ  $B^\alpha(L^p)$  μετρούν τη δύναμη των σημάτων στο πλάνο διαστήματος – κλίμακας: το  $L^p$  για μεταβολές ενδοκλίμακας  $L^q$  για μεταβολές δια- ή επι-κλίμακας (υπό την άποψη του  $d\lambda = dh/h$ ), ενώ το  $h^{-\alpha}$  για σύγκριση με τη συνέχεια του Χόλντερ.

Στον ορισμό, το  $h = \infty$  δεν προκαλεί κανένα πρόβλημα στην ακεραιότητα του (3.27) με την προϋπόθεση ότι τόσο το  $\alpha$  όσο και το  $q$  είναι θετικά. Το ευαίσθητο άκρο είναι το υπεριώδες όριο (δανειζόμενοι την οξυδερκή ορολογία των Μάμφορτ και Γκίτας [225] καθώς το «μήκος κύματος  $h \rightarrow 0$ ). Ας υποθέσουμε ότι για μια δεδομένη εικόνα  $u \in B_\alpha^\alpha(L^p)$ , το μέτρο  $P$  της συνέχειας έχει μια αυστηρή μείωση νόμου δύναμης στο υπεριώδες άκρο:

$$\omega_p(u, h) = O(h^\beta).$$

Τότε η μείωση πρέπει να είναι πιο γρήγορη από  $\alpha$ :  $\beta > \alpha$ , διαφορετικά, το υπεριώδες άκρο θα προκαλέσει έκρηξη στην ακεραιότητα (3.27). Είναι από αυτή τη θεώρηση που το  $\alpha$  αποκαλείται ο δείκτης κανονικότητας.

Είναι εύκολο να δούμε ότι το  $B_\infty^\alpha(L^\infty)$  είναι ακριβώς το διάστημα του Χόλντερ  $H^\alpha$ , το οποίο καθορίζεται συνήθως από μια μοναδική φόρμουλα

$$\sup_{x, y \in \mathbb{R}} \frac{|u(x) - u(y)|}{|x - y|^\alpha} < \infty.$$

Έτσι τα διαστήματα του Μπέσοβ γενικοποιούν φυσικά τα διαστήματα του Χόλντερ.

Αν το διάστημα κλίμακας  $h \in \mathbb{R}^+$  είναι προικισμένο με τη μέτρηση

$d\mu_{\alpha,q}(h) = h^{-1-\alpha q} dh$ , τότε η συνθήκη του Μπέσοβ γίνεται απλά

$\omega_p(u, \cdot) \in L^q(\mathbb{R}^+, d\mu_{\alpha,q})$ . Ή, αν κάποιος εισάγει το λογαριθμικό επίπεδο κλίμακας  $\lambda = -\ln h \in \mathbb{R}$  και γράψει  $\omega_p(u, \lambda) = \omega_p(u, h)$ , τότε η συνθήκη του Μπέσοβ γίνεται

$$\int_{\mathbb{R}} \omega_p(u, \lambda)^q e^{\lambda \alpha q} d\lambda < \infty,$$

από την οποία η υπερίωδης ευαισθησία όπως  $\lambda \rightarrow +\infty$  γίνεται επίσης εμφανής. Επιπλέον, παράλληλα με το MRA, ο φυσικό λογάριθμος μπορεί να αντικατασταθεί από ένα δυαδικό, και η συνθήκη του Μπέσοβ είναι ίση με

$$\int_{\mathbb{R}} \omega_p(u, \lambda)^q 2^{\lambda \alpha q} d\lambda < \infty. \quad (3.29)$$

Αν κάποιος διαχωρίσει επίσημα τον ακέραιο με  $d\lambda = 1$  στον πίνακα ακέραιων  $\mathbb{Z}$ , η συνθήκη Μπέσοβ γίνεται απλά

$$\sum \omega_p(u, j)^q 2^{j \alpha q} < \infty. \quad (3.30)$$

Κατα ακρίβεια, μια αυστηρή απόδειξη της ισότητας είναι εμφανής απλά παρατηρώντας ότι το  $\omega_p(u, h) = \omega_p(u, \lambda)$  αυξάνεται μονοτονικά όσον αφορά την κλίμακα  $h$ , ή ισοδύναμα,  $\lambda \in [j, j+1)$  μειώνεται όσον αφορά το  $\lambda$ . Συνεπώς, σε κάθε διάστημα

$$\omega_p(u, j+1)^q 2^{j \alpha q} \leq \omega_p(u, \lambda)^q 2^{\lambda \alpha q} \leq \omega_p(u, j)^q 2^{(j+1) \alpha q},$$

και ένα άθροισμα οδηγεί στο  $2^{-\alpha q} I_d \leq I \leq 2^{\alpha q} I_d$ , αν το  $I$  και το  $I_d$  υποδηλώνουν το ακέραιο και διακριτικό άθροισμα ξεχωριστά στις φόρμουλες (3.29) και (3.30).

Επιπλέον, εφόσον  $\omega_p(u, j) \leq 2 \|u\|_p$ , και  $\alpha, q > 0$ ,

$$\sum_{j < 0} \omega_p(u, j)^q 2^{j \alpha q} \leq \frac{2^q}{2^{\alpha q} - 1} \|u\|_p^q.$$

Αυτό εδραιώνει το ακόλουθο θεώρημα.

**Θεώρημα 3.9** Η μονάδα Μπέσοβ (3.28) είναι ίσο με

$$\|u\|_p + \left( \sum_{j \geq 0} 2^{j\alpha q} \omega_p(u, j)^q \right)^{1/q}.$$

Στο πιο πάνω έχει υποτεθεί ότι ο δείκτης κανονικότητας  $\alpha \in [0, 1)$ . Αν  $\alpha \in [n-1, n)$

για κάποιο θετικό ακέραιο  $n$ , ο ορισμός διαστήματος  $B_a^\alpha(L^p)$  του Μπέσοβ συνεχίζει να ισχύει με την προϋπόθεση ότι ο πεπερασμένος διαφοροποιητικός τελεστής πρώτης τάξης

$\Delta_a u = u(x+a) - u(x) = (T_a - I)u$ , όπου το  $T_a$  υποδηλώνει την αντίστροφη μετάφραση τελεστή και το  $I$  την ταυτότητα, αντικαθιστάται από τον πεπερασμένο διαφοροποιητικό τελεστή της τάξης

$$\Delta_a^n = (T_a - I)^n = \sum_{m=0}^n (-1)^{n-m} \binom{n}{m} T_{ma},$$

και το βήμα  $p$  συνέχειας  $\omega_p(u, h)$  στο (3.26) αναβαθμίζεται στο

$$\omega_{p,n}(u, h) = \sup_{|a| \leq h} \|\Delta_a^n u\|_p. \quad (3.31)$$

Για περισσότερες πληροφορίες σχετικά με τα διαστήματα Μπέσοβ παραπέμπουμε τον αναγνώστη στους Μέγιερ [216] και Γουοκτάσικ [324]. Για παράδειγμα, κάποιος θα μπορούσε ακόμα και να

επιτρέψει το  $\alpha$  να είναι αρνητικό από δυαδικότητα, και από την άλλη, με την προϋπόθεση ότι

$\alpha < n$ , ο ορισμός του  $B_q^\alpha(L^p)$  είναι ανεξάρτητος από το  $n$ .

### Χαρακτηρισμός Κυματιδίων στις Εικόνες Μπέσοβ

Δεδομένης μιας ορθογωνικής αλυσίδας MRA

$$\dots \subseteq V_{-1} \subseteq V_0 \subseteq V_1 \subseteq \dots,$$

Έστω ότι  $E_j$  και  $D_j$  υποδηλώνουν τις ορθογωνικές προβολές από το  $L^2(\mathbb{R})$  στο  $V_j$  και τα διαστήματα κυματιδίων  $W_j$  ξεχωριστά. Όπως και πριν, θα υποτεθεί ότι το μητρικό κυματίδιο  $\psi$  είναι  $r$ -κανονικό, και το  $r$  μπορεί να είναι οποιαδήποτε επιθυμητή τάξη.

Για οποιαδήποτε εικόνα  $u$ , έστω ότι  $u_j$  υποδηλώνει τη συνιστώσα κυματιδίου  $D_j u$  για οποιοδήποτε  $j \in \mathbb{Z}$ :

$$u_j(x) = \sum_k d_{j,k} \psi_{j,k}(x).$$

Λόγω της συνθήκης στιγμών εξαφάνισης στο (3.18), το μητρικό κυματίδιο  $\psi$  μοιάζει με τον τελεστή διαφοροποίησης  $d^{r+1}/dx^{r+1}$ . Επιπλέον, η προβολή  $D_j$  μοιάζει με τον πεπερασμένο τελεστή διαφοράς στην κλίμακα  $h_j = 2^{-j} : \Delta_{h_j}^{r+1}$ , γεγονός που δηλώνει ότι το βήμα  $P$  της συνέχειας στο (3.31),  $\omega_{p,r}(u, j)$ , είναι σχεδόν ίσο με το  $\|u_j\|_p$ . Ακολουθώντας το Θεώρημα 3.9, αυτό το βολικό (αλλά εριστικό) επιχείρημα οδηγεί σε ένα νέα χαρακτηρισμό των εικόνων του Μπέσοβ από την άποψη των συνιστώσων κυματιδίων τους:

$$\|E_0 u\|_p + \left( \sum_{j \geq 0} 2^{j\alpha q} \|u_j\|_p^q \right)^{1/q} < \infty. \quad (3.32)$$

Προσέξτε ότι το  $\|u\|_p$  στο Θεώρημα 3.9 έχει αντικατασταθεί από τη μονάδα  $P$  της βαθυπερατής προβολής  $E_0 u$ , η οποία φαίνεται πιο αδύναμη αλλά στην ουσία ισοδύναμη. Αυτό οφείλεται στο  $\|u_j\|_p = \|D_j u\|_p \leq C 2^{-j\alpha}$ , όπου το  $C$  υποδηλώνει ολόκληρο το δεύτερο όρο στο (3.32), και η σύγκλιση του

$$u = E_0 u + D_0 u + D_1 u + \dots$$

στο  $L^p$  είναι συνεπώς γεωμετρική και δυνατή δηλώνοντας ότι το  $u \in L^p$  και η μονάδα του  $\|u\|_p$  είναι όντως ελεγχόμενη από την αριστερή μεριά στο (3.32). Μια λεπτομερής αδιάσυστη απόδειξη για τους ισχυρισμούς αυτής της παραγράφου μπορεί να βρεθεί στο ωραίο βιβλίο [324].

Είμαστε σε καλή θέση να εισάγουμε το κύριο αποτέλεσμα αυτής της ενότητας, το οποίο είναι επίσης ένα γνωστό αποτέλεσμα στη θεωρία των κυματιδίων [215]. Με άλλα λόγια, οι μονάδες  $P$  των συνιστώσων  $\|u_j\|_p$  των κυματιδίων μπορούν να χαρακτηριστούν άμεσα από τους συντελεστές των κυματιδίων τους  $(d_{j,k} \mid k \in \mathbb{Z})$ .

**Θεώρημα 3.10 (Μονάδες Ισότητας Ανάμεσα σε  $L^p$  και  $l^p$ ).** Υπάρχουν δυο σταθερές  $C_1$  και  $C_2$ , και οι δυο εξαρτώνται μόνο από το μητρικό κυματίδιο  $\psi$ , τέτοιο που για κάθε ευκρίνεια  $j$ ,

$$C_1 \|u_j\|_p \leq 2^{j(1/2-1/p)} \|(d_{j,k} \mid k \in \mathbb{Z})\|_{l^p} \leq C_2 \|u_j\|_p. \quad (3.33)$$

Πρώτα, υποστηρίζουμε ότι αν η ισότητα ισχύει για την κλίμακα αναφοράς  $j = 0$ , πρέπει να ισχύει αυτόματα για όλα τα  $j$ . Για οποιαδήποτε  $j$ ,

$$u_j(x) = \sum_{k \in \mathbb{Z}} d_{j,k} 2^{j/2} \psi(2^j x - k).$$

Συνεπώς,  $v(x) = 2^{-j/2} u_j(2^{-j} x) \in W_0$  με συντελεστές κυματιδίων  $(d_{j,k} \mid k \in \mathbb{Z})$ . Εφαρμόζοντας την ισότητα στο  $v(x)$  για  $j = 0$ , έχουμε

$$C_1 \|2^{-j/2} u_j(2^{-j} x)\|_p \leq \|(d_{j,k} \mid k \in \mathbb{Z})\|_{l^p} \leq C_2 \|2^{-j/2} u_j(2^{-j} x)\|_p.$$

Ο ισχυρισμός συνεπώς επαληθεύεται εφόσον  $\|f(2^{-j} x)\|_p = 2^{j/p} \|f(x)\|_p$ .

(Επιπρόσθετα, αυτό το επιχείρημα δείχνει ότι αν ο τομέας εικόνας  $\Omega = \mathbb{R}^n$ , τότε ο παράγοντας  $2^{j(1/2-1/p)}$  βάρους θα έπρεπε να αντικατασταθεί από  $2^{jn(1/2-1/p)}$ .)

Συνεπώς, προκειμένου να εδραιώσουμε το Θεώρημα 3.10, αρκεί μόνο να αποδείξουμε το ακόλουθο γενικό αποτέλεσμα.

**Θεώρημα 3.11.** Ας υποθέσουμε ότι το  $w(x)$  είναι μια συνεχόμενη λειτουργία τέτοια που

1. όπως  $as x \rightarrow \infty, w(x) = O(|x|^{-\beta})$  για ορισμένα  $\beta > 1$ , και

2.  $(w(x - k) \mid k \in \mathbb{Z})$  είναι ορθογωνικά.

Τότε υπάρχουν δυο σταθερές  $C_1$  και  $C_2$ , που εξαρτώνται μόνο από το  $w(x)$  και  $p \geq 1$ , τέτοιες που για οποιαδήποτε λειτουργία  $f(x) = \sum_{k \in \mathbb{Z}} c(k) w(x - k)$

$$C_1 \|f\|_{L^p} \leq \|(c(k) \mid k \in \mathbb{Z})\|_{l^p} \leq C_2 \|f\|_{L^p}.$$

Για να αρχίσουμε, πρώτα αποδουκνύουμε ένα γενικό λήμμα.

**Λήμμα 3.12.** Έστω ότι  $d\mu_x$  και  $d\mu_y$  είναι δυο μετρήσεις στηριζόμενες στο  $x \in \mathbb{R}$  και  $y \in \mathbb{R}$ . Ας υποθέσουμε ότι το  $K(x, y)$  είναι μια μετρίσιμη λειτουργία στο διάστημα μέτρησης

προϊόντος  $(\mathbb{R}^2, d\mu_x \times d\mu_y)$ ,  $g(x) \in L^q(d\mu_x)$ , και  $c(y) \in L^p(d\mu_y)$

με  $p, q \geq 1$  οι οποίες κλίνουν η μια στην άλλη:  $p^{-1} + q^{-1} = 1$ . Τότε

$$\left| \int_{\mathbb{R}^2} K(x, y) c(y) g(x) d\mu_x d\mu_y \right| \leq \|L(x)\|_{\infty}^{1/q} \|R(y)\|_{\infty}^{1/p} \|g\|_q \|c\|_p, \quad (3.34)$$

όπου

$$L(x) = \int_{\mathbb{R}} |K(x, y)| d\mu_y \text{ and } R(y) = \int_{\mathbb{R}} |K(x, y)| d\mu_x, \quad (3.35)$$

και τα

δύο υποτιθέμενα ως πεπερασμένα.

Για την ακρίβεια, η ανισότητα (3.34) πηγάζει άμεσα από την ανισότητα του Χόλντερ,

$$\int_{\Omega} |F \cdot G| d\nu \leq \|F\|_{L^p(d\nu)} \|G\|_{L^q(d\nu)},$$

όπως εφαρμόζεται  $\Omega = \mathbb{R} \times \mathbb{R}$ ,  $F(x, y) = c(y)$ ,  $G(x, y) = g(x)$ , και η μέτρηση  $d\nu = |K| d\mu_x d\mu_y$ , του προϊόντος βάρους  $|K|$ .

Για να αποδείξουμε το Θεώρημα 3.11 από το λήμμα, ας προσέξουμε πρώτα ότι

$$c(k) = \int_{\mathbb{R}} w(x - k) f(x) dx, \quad k \in \mathbb{Z}.$$

Από την άλλη, εισάγοντας τη μέτρηση της σειράς- $\delta$

$$d\mu_y = \sum_{k \in \mathbb{Z}} \delta(y - k)$$

με το  $\delta$ , του Ντιρακ κάποιος έχει

$$f(x) = \int_{\mathbb{R}} w(x - y) c(y) d\mu_y,$$

όπου  $c(y)$  μπορεί να είναι οποιαδήποτε επέκταση του  $c(k)$  από το  $\mathbb{Z}$  στο  $\mathbb{R}$  καθώς το  $\mathbb{R} \setminus \mathbb{Z}$  είναι ένα μηδενικό σύνολο  $d\mu_y$ . Επιπλέον, σύμφωνα με το  $d\mu_y$  είναι σχεδόν βέβαιο ότι η εξίσωση  $f \rightarrow c$  μπορεί να επαναεκφραστεί ως

$$c(y) = \int_{\mathbb{R}} w(x - y) f(x) dx.$$

Έτσι εφαρμόζοντας το λήμμα στο  $K(x, y) = w(x - y)$ ,  $c(y) \in L^p(d\mu_y)$ , και οποιοδήποτε αυθαίρετο  $g(x) \in L^q(d\mu_x = dx)$ , κάποιος εξασφαλίζει

$$\left| \int_{\mathbb{R}} f(x)g(x)dx \right| \leq \|L\|_{\infty}^{1/q} \|R\|_{\infty}^{1/p} \|c\|_{l^p} \|g\|_q$$

το οποίο δηλώνει άμεσα ότι  $\|f\|_p \leq \|L\|_{\infty}^{1/q} \|R\|_{\infty}^{1/p} \|c\|_{l^p}$ . Εφόσον ο ρόλος των  $f$  και  $c$  είναι απόλυτα συμμετρικοί στο πιο πάνω επιχείρημα, η άλλη κατεύθυνση πρέπει να ισχύει

επίσης. Τέλος τόσο το  $\|L\|_{\infty}$  όσο και το  $\|R\|_{\infty}$  είναι όντως πεπερασμένα εφόσον

$$L(x) = \int_{\mathbb{R}} |K(x, y)| d\mu_y = \sum_k |w(x - k)|$$

είναι συνεχόμενο και περιοδικό με περίοδο 1,

$$R(y) = \int_{\mathbb{R}} |K(x, y)| d\mu_x = \int_{\mathbb{R}} |w(x - y)| dx \equiv \|w\|_1.$$

Για να φτάσουμε στο συμπέρασμα, ακολουθώντας το (3.32), μπορούσαμε να εδραιώσουμε τον αντίστοιχο χαρακτηρισμό μιας εικόνας Μπέσοβ  $u \in B_n^{\alpha}(L^p)$  με βάση τους συντελεστές κυματιδίων  $d_{j,k}$ , και τους συντελεστές βαθυπερατού  $c_{0,k}$  στην κλίμακα αναφοράς  $j = 0$ :

$$\|c_{0,\cdot}\|_{l^p} + \left( \sum_{j \geq 0} 2^{jq(\alpha+1/2-1/p)} \|d_{j,\cdot}\|_{l^p}^q \right)^{1/q} < \infty. \quad (3.36)$$

### Ανεπάρκεια Αναπαράστασης Κυματιδίων και Συμπίεσης Εικόνων

Οι εικόνες Μπέσοβ επιτρέπουν μια βολική ανάλυση της αναπαράστασης και υπολογισμού κυματιδίων. Παραπέμπουμε τον αναγνώστη στα, π.χ., αξιοσημείωτα γραπτά [55, 101, 102] για περισσότερες πληροφορίες.

Μια ιδιαίτερα ενδιαφέρουσα τάξη εικόνων του Μπέσοβ είναι  $B_p^{\alpha}(L^p)$ , της οποίας ο δείκτης ενδοκλίμακας  $p$  ισούται με το δείκτη διακλίμακας  $q$ , και η μονάδα Μπέσοβ μπορεί να οριστεί ως

$$\|u\|_{B_p^{\alpha}(L^p)} = \left( \sum_{k \in \mathbb{Z}} |c_{0,k}|^p \right)^{1/p} + \left( \sum_{j \geq 0, k \in \mathbb{Z}} 2^{jp(\alpha+1/2-1/p)} |d_{j,k}|^p \right)^{1/p}. \quad (3.37)$$

Συγκεκριμένα, όταν  $p = q$  και  $\alpha + 1/2 = 1/p$  (σε μια διάσταση), το υπολογισμένο άθροισμα στον συντελεστή κυματιδίου αποζηυγνύεται εντελώς κατά μήκος όλων των κλιμάκων:

$$\|u\|_{B_p^\alpha(L^p)} = \left( \sum_{k \in \mathbb{Z}} |c_{0,k}|^p \right)^{1/p} + \left( \sum_{j \geq 0, k \in \mathbb{Z}} |d_{j,k}|^p \right)^{1/p}. \quad (3.38)$$

Αυτά είναι τα μοντέλα εικόνων που μελετούνται συχνά στη λογοτεχνία της συμπίεσης εικόνων και υπολογισμού εικόνων (βλέπε π.χ, [55, 101, 102]).

Ως παράδειγμα, πρώτα ας θεωρήσουμε την αποτελεσματικότητα της γραμμικής συμπίεσης κυματιδίων. Ένα κλασσικό σχήμα γραμμικής συμπίεσης είναι η επιλογή ενός στοχευμένου επιπέδου ευκρίνειας  $J$ , τέτοιας που όλες οι συνιστώσες των κυματιδίων που είναι καλύτερες από το  $J$  παρατούνται. Με άλλα λόγια, η συμπίεσμένη εικόνα  $u_J$  για κάθε δεδομένη εικόνα  $u$  καθορίζεται ως

$$u_J = E_0 u + \sum_{0 \leq j \leq J} D_j u.$$

Το λάθος συμπίεσης, μετρημένο στο μετρικό  $L^2$ , δίνεται από το

$$\|u - u_J\|_2^2 = \sum_{j > J} \|D_j u\|_2^2 = \sum_{j > J, k \in \mathbb{Z}} |d_{j,k}|^2.$$

Ακολουθώντας το [55], ας υποθέσουμε ότι  $u$  είναι μια εικόνα Μπέσοβ  $B_2^\alpha(L^2)$  για κάποιο  $\alpha > 0$ , το οποίο σύμφωνα με το (3.37) δηλώνει

$$\sum_{j \geq 0, k \in \mathbb{Z}} 2^{2j\alpha} |d_{j,k}|^2 \leq \|u\|_{B_2^\alpha(L^2)}^2.$$

Συνεπώς, το σφάλμα  $L^2$  για ένα τέτοιο σχήμα γραμμικής συμπίεσης ελέγχεται από

$$\begin{aligned} \|u - u_J\|_2^2 &= \sum_{j > J, k \in \mathbb{Z}} |d_{j,k}|^2 \\ &\leq 2^{-2J\alpha} \sum_{j > J, k \in \mathbb{Z}} 2^{2j\alpha} |d_{j,k}|^2 \\ &\leq 2^{-2J\alpha} \|u\|_{B_2^\alpha(L^2)}^2, \end{aligned}$$

ή απλά,



$$\|u - u_J\|_2 \leq 2^{-J\alpha} \|u\|_{B_2^\alpha(L^2)} = h_J^\alpha \|u\|_{B_2^\alpha(L^2)},$$

όπου το  $h_J = 2^{-J}$  είναι η χωρική ευκρίνεια η οποία είναι καλύτερη από αυτή στην οποία είναι ρυθμισμένοι στο 0 όλοι οι συντελεστές κυματιδίων. Είναι εμφανές από αυτό το παράδειγμα ότι η αποτελεσματικότητα του σχήματος γραμμικής συμπίεσης είναι ανάλογη της κανονικότητας της δεδομένης εικόνας (π.χ.,  $\alpha$ ).

Ακολουθώντας το [55, 101] ας θεωρήσουμε τώρα ένα δημοφιλές σχήμα μη-γραμμικής συμπίεσης που εφαρμόζεται διατηρώντας μόνο τους μεγαλύτερους συντελεστές κυματιδίων  $N$ .

Ας υποθέσουμε ότι το  $u$  είναι μια εικόνα του Μπέσοβ που ανήκει στο  $B_p^\alpha(L^p)$  με  $\alpha + 1/2 = 1/p$  (σε μια διάσταση, ή  $\alpha/2 + 1/2 = 1/p$  σε δυο διαστάσεις· είτε η περίπτωση δηλώνει  $p < 2$ ). Τότε από το (3.38),

$$\sum_{k \in \mathbb{Z}} |c_{0,k}|^p + \sum_{j \geq 0, k \in \mathbb{Z}} |d_{j,k}|^p \leq \|u\|_{B_p^\alpha(L^p)}^p.$$

Το σχήμα μη γραμμικής συμπίεσης αρχίζει με την επαναδιάταξη όλων των συντελεστών  $(c_{0,k} \mid k \in \mathbb{Z})$  και  $(d_{j,k} \mid j \geq 0, k \in \mathbb{Z})$  από τη φθίνουσα σειρά των απόλυτων αξιών:

$$a_1, a_2, \dots, a_{N-1}, a_N, a_{N+1}, \dots \text{ με } |a_1| \geq |a_2| \geq |a_3| \geq \dots.$$

Η συμπίεσμένη εικόνα ορίζεται ως

$$u_N(x) = \sum_{n=1}^N a_n g_n(x),$$

όπου το  $g_n$  υποδηλώνει είτε  $\phi_{0,k}$  ή  $\psi_{j,k}$  αν  $a_n = c_{0,k}$  or  $d_{j,k}$ .

Για να υπολογίσουμε το σφάλμα προσέγγισης ενός τέτοιου σχήματος μη γραμμικής συμπίεσης, πρώτα γράφουμε  $\lambda_N = |a_N|$ , και προσέχουμε

$$N \lambda_N^p \leq \sum_{n=1}^{\infty} |a_n|^p = \sum_{k \in \mathbb{Z}} |c_{0,k}|^p + \sum_{j \geq 0, k \in \mathbb{Z}} |d_{j,k}|^p \leq \|u\|_{B_p^\alpha(L^p)}^p,$$

οποίο δηλώνει ότι

το

$$\lambda_N^p \leq N^{-1} \|u\|_{B_p^\alpha(L^p)}^p.$$

Συνεπώς προσέχοντας ότι  $\alpha = 2/p - 1$ , έχουμε

$$\begin{aligned} \|u - u_N\|_2^2 &\leq \sum_{n>N} |a_n|^2 \leq \lambda_N^{2-p} \sum_{n>N} |a_n|^p \\ &\leq \lambda_N^{2-p} \sum_{n=1:\infty} |a_n|^p \\ &\leq (\lambda_N^p)^{(2/p-1)} \|u\|_{B_p^\alpha(L^p)}^p \\ &\leq N^{-\alpha} \|u\|_{B_p^\alpha(L^p)}^2, \end{aligned}$$

ή απλά,

$$\|u - u_N\|_2 \leq N^{-\alpha/2} \|u\|_{B_p^\alpha(L^p)} = O(N^{-\alpha/2}). \quad (3.39)$$

Όπως και στην περίπτωση της γραμμικής συμπίεσης το σφάλμα προσέγγισης ελέγχεται άμεσα από το δείκτη ομαλότητας  $\alpha$ .

Επιπλέον, όπως παρουσιάζεται από τους Ντεβόρ, Τζάγουερθ, και Πόποβ [102], η τελευταία φόρμουλα ελέγχου σφάλματος είναι αντιστρέψιμη. Γενικά μιλώντας, σημαίνει ότι αν για οποιαδήποτε δεδομένη

εικόνα  $u$  υπάρχει μια σταθερά  $C > 0$  και  $\alpha > 0$ , τέτοια που το πιο πάνω σχήμα μη-γραμμικής συμπίεσης παράγει

$$\|u - u_N\|_2 \leq CN^{-\alpha/2},$$

τότε το  $u \in B_p^\alpha(L^p)$  για  $p$  με  $1/p = \alpha + 1/2$  (σε μια διάσταση ή

$1/p = \alpha/2 + 1/2$  σε δυο διαστάσεις).

Στο επόμενο κεφάλαιο, θα εξηγήσουμε το ρόλο των εικόνων Μπέσοβ στην αποθορυβοποίηση εικόνων.

### 3.4 Πλεγματική και Τυχαία Αναπαράσταση Πεδίου

Μακριά από όλες τις προηγούμενες πιο προσδιοριστικές προσεγγίσεις στη μοντελοποίηση εικόνων, σε αυτή την ενότητα σχολιάζουμε στοχαστικές μεθόδους βασισμένες στα τυχαία πεδία των Γκιπς και Μάρκοβ [130, 328, 329].

#### 3.4.1 Φυσικές Εικόνες της Μητέρας Φύσης

Κοιτώντας μια τυπική φωτογραφία που πάρθηκε κατά τη διάρκεια ενός ταξιδιού διακοπών στην παραλία, είμαστε συχνά σε θέση να αναγνωρίσουμε αμέσως τον ωκεανό, τον μπλε ουρανό, την άμμο, τους φοίνικες, και λοιπά. Οι παραλίες δεν είναι πανομοιότυπες σε δυο διαφορετικές πόλεις, λόγω της διαφορετικότητας στο μέγεθος, το χρώμα, τον τύπο κόκκου και την αγνότητα της άμμου. Παρόλα αυτά, ένα κανονικό σύστημα ανθρώπινης όρασης δεν το βρίσκει ποτέ πολύ δύσκολο να τα αναγνωρίσει.

Προφανώς, το σύστημα όρασης πρέπει να μπορεί να μετρήσει και να αναλύσει συγκεκριμένες στατιστικά αμετάβλητες ιδιότητες προκειμένου να διατηρηθεί τέτοια δύναμη, η οποία δείχνει προς τη σημαντικότητα της κατάλληλης μοντελοποίησης των στοχαστικών χαρακτηριστικών φυσικών εικόνων.

Η λέξη «φυσικό» έχει χρησιμοποιηθεί ευρέως στην ανάλυση εικόνας και όρασης [122, 123, 142, 225]. Παρόλο που υστερεί σε ένα ομοιόμορφο προσδιορισμό, διαισθητικά αναφέρεται σε εικόνες που παγιδεύουν τυχαία σκηνές στη φύση, όπως τις αμμώδεις παραλίες που αναφέρθηκαν πιο πάνω ή πράσινα γρασιδία. Οι φυσικές εικόνες αποκαλούνται συχνά και «υφές». Σε αντίθεση με τις «τεχνητές» εικόνες που παγιδεύουν κυρίως ομαλές σκηνές δημιουργημένες από τον άνθρωπο, οι υφές είναι συχνά ακατέργαστες και διακυμαίνονται όσον αφορά τις λειτουργίες εικόνων.

Το ακατέργαστο και η διακύμανση στις φυσικές εικόνες είναι άμεσο αποτέλεσμα από εκείνα της Μητέρας Φύσης. Αυτά τα χαρακτηριστικά είναι συχνά μη-άμεσα ζωγραφισμένα σε δυσδιάστατες επιφάνειες της Μητέρας Φύσης· αντί για αυτό είναι χαραγμένα στον τρισδιάστατο κόσμο από πολυάριθμες τυχαίες φυσικές δυνάμεις, όπως κυλιόμενες παλίρροιες, το πλύσιμο της βροχής, την κίνηση του ανέμου και βιολογικές καλλιέργειες ελεγχόμενες από τον καιρό. Η επακόλουθη τρισδιάστατη τραχύτητα και διακύμανση στις επιφάνειες γεωμετρίας παγιδεύονται έπειτα από τις δυσδιάστατες εικόνες και γίνονται δυσδιάστατες υφές. Οι αμμώδεις παραλίες, τα δέντρα, λειμώνες και λοφώδεις σκηνές είναι όλα τέτοια παραδείγματα.

Υπάρχουν ορισμένες προσδοκίες στις υφές που είναι βασισμένες στην τραχύτητα (ή στη γεωμετρία επιφάνειας). Εκείνα είναι υπό κάποια έννοια πραγματικά ζωγραφισμένα από τη Μητέρα Φύση σε σχετικά ομαλότερες επιφάνειες. Οικεία παραδείγματα περιλαμβάνουν στην ξυλινη υφή σε μια ομαλά κομμένη περιοχή ενός κορμού και πολυάριθμων μοτίβων δερμάτων ζώων (π.χ. ζέβρες, λεοπάρδαλες, φίδια) δημιουργημένα από βιοχημικούς μηχανισμούς (π.χ. τα μοντέλα αντίδρασης – διάχυσης του Τούριγκ [304]).

Συμπερασματικά, οι δυσδιάστατες υφές είναι στην ουσία οι καθρεπτισμοί της Μητέρας Φύσης, με πιστούς αντικατοπτρισμούς της τυχειότητας και των διακυμάνσεών Της. Συνεπώς, τα στοχαστικά εργαλεία είναι ενδογενώς απαραίτητα για την κατάλληλη αναπαράσταση και ανάλυση φυσικών εικόνων.

### **3.4.2 Εικόνες σαν Σύνολα και Κατανομές**

Σε αντίθεση με την προσδιοριστική άποψη σύμφωνα με την οποία κάθε εικόνα είναι μια εικόνα από μόνη της, η στατιστική άποψη αντιμετωπίζει κάθε εικόνα σαν ένα δείγμα από το σύνολο στο οποίο ανήκει.

Έτσι, για παράδειγμα, παρόλο που μια εικόνα γρασιδιού που πάρθηκε στο Παρίσι ίσως να φαίνεται σε κάθε λεπτομέρεια πολύ διαφορετική από εκείνη που πάρθηκε στη Νέα Υόρκη, εξακολουθούν να είναι εύκολα αναγνωρίσιμες σαν εικόνες γρασιδιού, ή δυο διαφορετικά δείγματα από το σύνολο των εικόνων που είναι κοινώς επώνυμες από τη λέξη «γρασιδί».

Από την άλλη, η ανθρώπινη όραση είναι επίσης αρκετά ανεπτυγμένη για να ξεχωρίσει μια τυπική εικόνα αμμώδους παραλίας από μια τυπική εικόνα γρασιδιού. Με άλλα λόγια, έχει την αξιόλογη ικανότητα να ταξινομεί διαφορετικά σύνολα εικόνων. Ακόμη και στο επίπεδο συνόλου, σε αντίθεση με την προσδιοριστική άποψη, διαφορετικά σύνολα εικόνων πρέπει να μεταφέρουν αχρείαστα ξεκάθαρα όρια αποφάσεων. Υπάρχουν συγκεκριμένα δείγματα εικόνων που ίσως φαίνονται τόσο σαν εικόνες γρασιδιού όσο και σαν εικόνες αμμώδων παραλιών. Για παράδειγμα, τυπικά κοντά σε μια άστρωτη και

μη-ανεπτυγμένη φυσική παραλιακή περιοχή, καθώς περπατά κάποιος μακριά από την ακτή στη γη, η αγνή αμμώδης παραλία μπορεί σταδιακά να μεταφέρει περισσότερο έδαφος και να επιτρέψει τη βλάστηση περισσότερου γρασιδιού. Ως αποτέλεσμα, από άποψη των εικόνων τους, τα αμμώδη χαρακτηριστικά ίσως εξαφανίζονται αργά καθώς εκείνα του γρασιδιού γίνονται εντονότερα. Σε ένα συγκεκριμένο σημείο, ακόμη και στην ανθρώπινη όραση μια δυαδική απόφαση άμμου ή γρασιδιού γίνεται πολύ δύσκολη.

Έτσι, αν το  $\mathcal{I} = \mathcal{I}(\Omega)$  υποδηλώνει όλες τις πιθανές εικόνες σε ένα δυδιάστατο τομέα  $\Omega$ , η στοχαστική άποψη δεν σκοπεύει να τμηματοποιήσει το  $\mathcal{I}$  σε κομματιασμένα ξεχωριστά υποσύνολα όπως

$$\mathcal{I} = \mathcal{I}_{\text{grass}} \cup \mathcal{I}_{\text{tree}} \cup \mathcal{I}_{\text{sand beach}} \cup \dots$$

με κάθε συλλογή να αναπαριστά αποκλειστικά μια συγκεκριμένη κατηγορία.

Αντιθέτως, οποιαδήποτε εικόνα στο  $\mathcal{I}$  πιστεύεται ότι έχει μια συγκεκριμένη πιθανότητα να χαρακτηριστεί σαν εικόνα γρασιδιού, ακόμη και αν μοιάζει «εντελώς» με δέντρο στην ανθρώπινη όραση και αντίστροφα. Μαθηματικά, αυτό δηλώνει ότι κάθε σύνολο είναι στην ουσία μια πιθανότητα κατανομής  $\mu$  στο  $\mathcal{I}$ , και διαφορετικές κατηγορίες αναλογούν σε διαφορετικές κατανομές. Για παράδειγμα, ίσως έχουμε  $\mu_{\text{grass}}$ ,  $\mu_{\text{tree}}$ , και  $\mu_{\text{sandbeach}}$  για τα σύνολα του γρασιδιού, του δέντρου και της αμμώδους παραλίας.

Συνεπώς, η αποστολή της στοχαστικής αναπαράστασης και ανάλυσης φυσικών εικόνων είναι να αναπτύξουμε και να μοντελοποιήσουμε κατάλληλα αυτές τις κατανομές συνόλων  $\mu$ . Έχει ευρέες εφαρμογές στις σύγχρονες επιστήμες και τεχνολογίες, για παράδειγμα, ανάγνωση εδάφους για αυτόματο πιλοτάρισμα, ανίχνευση στόχου και αναγνώριση στο στρατό και αυτόματη ανίχνευση ακανόνιστων ιστών στην εικονική ιατρική, για να ονομάσουμε μόνο μερικές.

Αυτά που αποκαλούνται στην καθημερινή ζωή «χαρακτηριστικά» ή «μοτίβα» εικόνων αναλογούν στα χαρακτηριστικά τέτοιων κατανομών συνόλων  $\mu$ .

Αυτή η στοχαστική άποψη έχει υπάρξει για την ακρίβεια πολύ ουσιαστική για την θεωρία της κλασσικής ανίχνευσης σημάτων. Πιο κάτω αναλύουμε ένα απλό παράδειγμα.

Ας θεωρήσουμε την ανίχνευση σημάτων δυο σημείων: ένα πιο φωτεινό και το άλλο πιο σκοτεινό. Στα αληθινά πειράματα, ίσως δημιουργηθούν ελέγχοντας τον φωτισμό της πηγής φωτός στο σκοτεινό δωμάτιο. Λόγω ηλεκτρικών ή περιβαλλοντικών διακυμάνσεων, οι εντάσεις φωτός των δυο σημείων διακυμαίνονται επίσης και μοντελοποιούνται τυπικά από τις Γκαουσιανές κατανομές, για παράδειγμα,

$$p^b(i) = \frac{1}{\sqrt{2\pi\sigma_b^2}} e^{-\frac{(i-4)^2}{2\sigma_b^2}} \quad \text{και} \quad p^d(i) = \frac{1}{\sqrt{2\pi\sigma_d^2}} e^{-\frac{(i-2)^2}{2\sigma_d^2}}. \quad (3.40)$$

Το φωτεινό σημείο έχει μέση ένταση  $m_b = 4$  και ρυθμό επέκτασης (ή ασάφεια)  $\sigma_b$ , ενώ το πιο σκοτεινό έχει μέση ένταση  $m_d = 2$  και ρυθμό επέκτασης  $\sigma_d$ . Αυτή είναι η στοχαστική άποψη των

δύο φαινομενικά προσδιοριστικών σημείων. Όταν  $\sigma_b = \sigma_d = 0$ , αυτά είναι δύο ιδανικά προσδιοριστικά σημεία.

Ας υποθέσουμε τώρα ότι  $\sigma_b = \sigma_d = 1$ , και μια συγκεκριμένη μέτρηση δίνει  $i_0 = 3$ . Τότε χωρίς επηρεασμό από οποιουδήποτε άλλους παράγοντες, ένας πειραματιστής θα έχει δυσκολία να καθορίσει ποιο σημείο ήταν παρόν. Αυτή είναι η ασάφεια στην ανίχνευση μοτίβων σημάτων, ακριβώς ανάλογα στην αναγνώριση μοτίβων: οι στηρίξεις των δυο κατανεμιστικών πιθανοτήτων μπορούν να αλληλοκαλύπτονται.

Σε αυτό το παράδειγμα, τα μοτίβα σήματος «φωτεινό» και «σκοτεινό» χαρακτηρίζονται στην πραγματικότητα από το ζευγάρι χαρακτηριστικών (μέσο, μεταβλητή)  $= (m, \sigma)$  στο Γκαουσιανό μοντέλο (3.40). ως αποτέλεσμα, μια ολοκληρωμένη και αποτελεσματική αναπαράσταση ενός σημείου δεν είναι η καταγραφή ξεχωριστών δειγμάτων  $i$ , αλλά μόνο η αποθήκευση των δυο χαρακτηριστικών αριθμών κλειδί: μέσο και μεταβλητή. Αυτή είναι και πάλι ανάλογη στο τι χρειάζεται να γίνει για τη στοχαστική αναπαράσταση και ανάλυση εικόνας: να εξάγουμε το μικρότερο ποσοστό στατιστικών πληροφοριών κλειδί έτσι ώστε το στοχευμένο σύνολο εικόνων να είναι χαρακτηρισμένο αποτελεσματικά.

Πιο κάτω εισάγουμε τα αξιολογούμενα έργα των Γκέμαν και Γκέμαν [130], Ζου και Μάμφορτ [328], και Ζου, Γου και Μάμφορτ [329] στη στοχαστική μοντελοποίηση εικονικών μοτίβων.

### 3.4.3 Εικόνες σαν Σύνολα Γκιπς.

Στις πλείστες ψηφιακές εφαρμογές, η εικονοποίηση ή έκθεση συσκευών αποτελούνται συχνά από δυσδιάστατες αλληλουχίες μικτροστοιχείων, τα περισσότερα από τα οποία είναι Καρτσιανά. Μετά την επανακλιμάκωση, έστω ότι το ακέραιο πλέγμα

$$\mathbb{Z}^2 = \{(n, m) : n, m = 0, \pm 1, \pm 2, \dots\}$$

υποδηλώνει την κανονική αλληλουχία εικόνας. Τότε ένας πεπερασμένος τομέας  $\Omega$  είναι ένα υποσύνολο του  $\mathbb{Z}^2$ , και τυπικά ένα τετράγωνο στις πλείστες ψηφιακές εφαρμογές:

$$\Omega = \{(i, j) : 0 \leq i \leq I - 1 \text{ and } 0 \leq j \leq J - 1\},$$

οι οποίες θα θεωρηθούν πιο κάτω εκτός και αν δηλωθεί διαφορετικά.

Όπως και πριν έστω ότι  $\mathcal{I}(\Omega)$  υποδηλώνουν όλες τις ψηφιακές εικόνες στο  $\Omega$ :

$$\mathcal{I}(\Omega) = \{u = (u_{ij}) : \Omega \rightarrow \mathbb{R} \mid u_{ij} \in \mathbb{R}\}.$$

Επίσημα, αυτά είναι απλά όλα τα  $I$ -από πίνακες  $J$  με πραγματικές καταχωρήσεις, και συνεπώς  $\mathcal{I} = \mathbb{R}^{I \times J}$ . Κάποιος, παρόλα αυτά πρέπει να συνειδητοποιήσει ότι η επίσημη δομή πίνακα μιας ψηφιακής εικόνας εφαρμόζεται σπάνια με την έννοια της γραμμικής άλγεβρας.

Η αποστολή της στοχαστικής μοντελοποίησης μιας τάξης εικόνων είναι να μοντελοποιήσει την υποκείμενη κατανομιστική πιθανότητα  $P$  στο  $\mathcal{I}(\Omega)$ . Προφανώς,  $P_{\text{grass}}$  θα έπρεπε να είναι λογικά διαφορετικό από το  $P_{\text{sand beach}}$ . Μια οικεία κατανομή στο  $\mathcal{I}(\Omega)$  είναι η ομοιογενής Γκαουσιανή

$$p(u) = \frac{1}{(2\pi\sigma^2)^{IJ/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{ij \in \Omega} u_{ij}^2\right) = \frac{1}{(2\pi\sigma^2)^{IJ/2}} \exp\left(-\frac{\|u\|^2}{2\sigma^2}\right), \quad (3.41)$$

όπου το  $\|\cdot\|$  αντιπροσωπεύει την Ευκλείδιο μονάδα στο  $\mathcal{I}$  ως το Ευκλείδιο διάστημα  $\mathbb{R}^{I \times J}$ . Δυστυχώς, ένα τυπικό δείγμα από αυτή την κατανομή φαίνεται συνήθως θορυβώδες και μη-ενδιαφέρον.

Παρόλα αυτά, υπάρχουν περισσότερα χαρακτηριστικά της Γκαουσιανής έκφρασης που είναι στη ουσία διαμοιρασμένα από πολλές άλλες κατανομές εικόνας. Ας ορίσουμε

$$E[u] = \frac{\|u\|^2}{2}, \quad \beta = \frac{1}{\sigma^2}, \quad \text{and} \quad Z = (2\pi\sigma^2)^{\frac{IJ}{2}}.$$

Τότε η Γκαουσιανή κατανομή θα μπορούσε να εκφραστεί ως

$$p(u) = p_{\beta}(u) = \frac{1}{Z} \exp(-\beta E[u]). \quad (3.42)$$

Αποκαλώντας επίσημα το  $E[u]$  την «ενέργεια» του  $u$ , έχουμε εξασφαλίσει τον ΕΣ του Γκιπς στη στατιστική μηχανική (βλέπε προηγούμενο κεφάλαιο).

Στη στατιστική μηχανική, το  $u$  υποδηλώνει μια μικροσκοπική κατάσταση ενός στοχευμένου συνόλου κυματιδίων και το  $E[u]$  τα επίπεδα ενέργειας εκείνης της συγκεκριμένης κατάστασης. Αν ένα δυσδιάστατο πλέγμα σωματιδίων όλα ελεύθερα να κινούνται προς μια συγκεκριμένη κατεύθυνση και το  $u_{ij}$  είναι συνδεδεμένο με την ταχύτητα του σωματιδίου που χαρακτηρίζεται από το  $(ij)$ , τότε η Γκαουσιανή κατανομή (3.41) γίνεται ο φημισμένος νόμος κατανομής Μάξγουελ-Μπόλτ兹μαν που υποθέτει ότι η μάζα κάθε κυματιδίου είναι μια μονάδα [82, 131].

Από την άλλη, στη στατιστική μηχανική η παράμετρος  $\beta$  αποκαλείται η «αντίστροφη» θερμοκρασία όπως δίνεται αποκλειστικά από το  $\beta = 1/(kT)$ , με το  $k$  να υποδηλώνει τη σταθερά Μπόλτ兹μαν και το  $T$  την απόλυτη θερμοκρασία.

Τέλος, η σταθερά  $Z$  αποκαλείται η λειτουργία τμήματος και είναι μοναδικά προσδιορισμένη από το  $\beta$  για πιθανότητα κανονικοποίησης

$$Z = Z(\beta) = \sum_{u \in \mathcal{I}} \exp(-\beta E[u]) \text{ or } \int_{u \in \mathcal{I}} \exp(-\beta E[u]) du,$$

όταν η ενέργεια έχει καθοριστεί. Εδώ το άθροισμα εφαρμόζεται όταν το  $\mathcal{I}$  είναι ποσοτικοποιημένο σε πεπερασμένα επίπεδα γκριζας κλίμακας και το  $\mathcal{I}$  είναι ένα πεπερασμένο σύνολο, ενώ η ακεραιότητα

όταν τα  $\mathbf{u}$  είναι συνεχόμενα και το  $\mathcal{I}$  είναι ένα ανοιχτό υποσύνολο στο  $\mathbb{R}^{I \times J}$ . Φυσικά, η λειτουργία τμήματος γενικά περιλαμβάνει σημαντικές πληροφορίες ελεύθερων ενεργειών και διακυμάνσεων [82, 131].

Για τη μοντελοποίηση εικόνων συχνά δεν αρκεί να δουλεύουμε απλά με μια μοναδική ενέργεια.

**Ορισμός 3.13 (Μοντέλα Εικόνων Γκιπς (Γκέμαν και Γκέμαν [130])).** Ένα πεδίο Γκιπς ή ένα σύνολο μοντέλου μιας τάξης εικόνας αναφέρεται σε οποιαδήποτε κατανομιστική πιθανότητα  $p(\mathbf{u})$  στο διάστημα εικόνας  $\mathcal{I}$ , το οποίο δίνεται από την ακόλουθη μορφή:

$$p(\mathbf{u}) = p_{\beta}(\mathbf{u}) = \frac{1}{Z} \exp(-\beta \cdot \mathbf{E}[\mathbf{u}]), \quad (3.43)$$

όπου για κάποιο πεπερασμένο  $m$ ,

$$\beta = (\beta_1, \beta_2, \dots, \beta_m) \text{ and } \mathbf{E}[\mathbf{u}] = (E_1[\mathbf{u}], E_2[\mathbf{u}], \dots, E_m[\mathbf{u}]) \quad (3.44)$$

είναι ορισμένοι από τους κατάλληλους παραμετρικούς φορείς και τον φορέα λειτουργίας «ενέργειας» ξεχωριστά, και η λειτουργία τμήματος δίνεται από

$$Z = Z(\beta) = \sum_{u \in \mathcal{I}} \exp(-\beta \cdot \mathbf{E}[\mathbf{u}]).$$

Όπως στη στατιστική μηχανική και το προηγούμενο κεφάλαιο, θα αποκαλούμε κάθε  $E_k[\mathbf{u}]$  μια γενικοποιημένη ενέργεια και τη δυαδική του παράμετρο  $\beta_k$  μια γενικοποιημένη προοπτική. Στη θερμοδυναμική για παράδειγμα, κάποιος μπορεί να έχει

$$\mathbf{E} = (E, V, n),$$

με την εσωτερική ενέργεια  $E$ , την ένταση  $V$ , και το μοριακό αριθμό  $n$  των σωματιδίων και

$$\beta = (\beta, p, \mu)$$

που αναλογεί στην αντόστροφη θερμοκρασία  $\beta$ , την πίεση  $p$ , και τη χημική προοπτική  $\mu$ .

Η κύρια πρόκληση για τη μοντελοποίηση εικόνας του Γκιπς είναι ο κατάλληλος σχεδιασμός αυτών των γενικοποιημένων λειτουργιών ενέργειας και των δυαδικών τους προοπτικών.

### 3.4.4 Οι Εικόνες ως Τυχαία Πεδία του Μάρκοβ

Υπάρχει μια ξεχωριστή τάξη ενεργειών των οποίων τα συνδεδεμένα σύνολα Γκιπς είναι τα τυχαία πεδία του Μάρκοβ. Το αντίθετο είναι επίσης αλήθεια, αποτέλεσμα αξιοσημείωτο στη στοχαστική μοντελοποίηση και ανάλυση λόγω των Χάμερсли και Κλιφορτ (βλέπε π.χ., [33]).

Όπως και πριν, έστω ότι  $\Omega$  υποδηλώνει ένα ιδιαίτερο σύνολο μικροστοιχείων, πάνω στα οποία προσδιορίζεται μια ένταση εικόνας  $u : \Omega \rightarrow \mathbb{R}$ . Ας υποθέσουμε ότι συνδεδεμένη με κάθε μικροστοιχείο  $\alpha \in \Omega$ , υπάρχει μια μοναδική γειτονιά  $N_\alpha \subseteq \Omega \setminus \{\alpha\}$ . Για παράδειγμα, για τον κανονικό Καρτεσιανό πίνακα  $\Omega = \mathbb{Z}^2$ , μια οικεία επιλογή είναι να προσδιορίσουμε για κάθε μικροστοιχείο  $\alpha = (i, j)$ ,

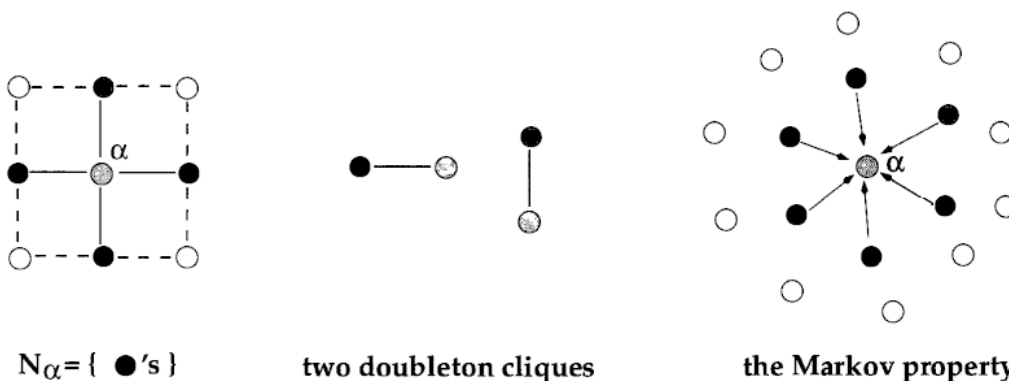
$$N_\alpha = \{\beta = (k, l) \in \Omega : |k - i| + |l - j| = 1\} = \{(i, j \pm 1), (i \pm 1, j)\}. \quad (3.45)$$

Προσέξτε ότι το  $\alpha$  εξαιρείται από τη γειτονιά του. Τότε η συλλογή όλων των γειτονιών

$$\mathcal{N} = \{N_\alpha \mid \alpha \in \Omega\}$$

αποκαλείται ένα γειτονικό σύστημα, το οποίο πάντα θα θεωρείται συμμετρικό με την έννοια ότι

$$\beta \in N_\alpha \text{ μόνο και μόνο αν } \alpha \in N_\beta, \quad \alpha, \beta \in \Omega.$$



**Εικόνα 3.4** Συστατικά των τυχαίων πεδίων του Μάρκοβ μέσω παραδειγμάτων: μια γειτονιά  $N_\alpha$

(αριστερά), δυο διπλές κλίκες  $C \in \mathcal{C}$  (μέση), και τοπικότητα της συμπερασματικής συνθήκης  $p(u_\alpha \mid u_{\Omega \setminus \alpha}) = p(u_\alpha \mid u_{N_\alpha})$  (δεξιά).



Τότε το ζευγάρι  $(\Omega, \mathcal{N})$  λέγεται ότι είναι ένα γράφημα ή τοπολογικός τομέας μικροστοιχείου. Θα υποδηλώσουμε επίσης  $\beta \in N_\alpha$  by  $\beta \sim \alpha$  ή  $\beta \sim_{\mathcal{N}} \alpha$ . Στη θεωρία γραφήματος, το γειτονικό σύστημα  $\mathcal{N}$  είναι επίσης ίσο με το σύστημα άκρου  $E_\Omega \subseteq \Omega \times \Omega$  με

$$E_\Omega = \{(\alpha, \beta) \in \Omega \times \Omega \mid \beta \sim \alpha\}.$$

Για απλότητα στην εγγραφή, το ίδιο σύμβολο  $u$  θα υποδηλώνει τόσο το τυχαίο πεδίο εικόνας όσο και οποιοδήποτε από τα ιδιαίτερα του δείγματα.

**Ορισμός 3.14 (Ιδιότητα του Μάρκοβ).** Ένα τυχαίο πεδίο  $u$  στον τομέα μικροστοιχείου  $\Omega$  λέγεται να είναι Μαρκοβιανό σε σχέση με το γειτονικό σύστημα  $\mathcal{N}$  αν είναι τοπικό, π.χ., όσων αφορά τις υποθετικές πιθανότητες, για οποιοδήποτε μικροστοιχείο  $\alpha \in \Omega$ ,

$$p(u_\alpha \mid u_{\Omega \setminus \{\alpha\}}) = p(u_\alpha \mid u_{N_\alpha}).$$

Με άλλα λόγια, οι πληροφορίες εικόνας σε οποιοδήποτε δεδομένο μικροστοιχείο  $\alpha$ , δεδομένης της υπόλοιπης εικόνας, εξαρτάται για την ακρίβεια μόνο από το γειτονικό  $N_\alpha$  (βλέπε Εικόνα 3.4).

Η τοπικότητα είναι επιθυμητή τόσο για τη θεωρητική ανάλυση όσο και για τον αποτελεσματικό υπολογισμό (όπως η παραλληλοποίηση), και τα τυχαία πεδία του Μάρκοβ έχουν συνεπώς παίξει σημαντικούς ρόλους στην μελέτη χωρικών μοτίβων κατά μήκος ενός αριθμού πειθαρχιών.

Για να δηλώσουμε το θεώρημα ισότητας των Γκιπς – Μάρκοβ που αναφέρθηκε στην αρχή είναι αναγκαίο να εισάγουμε την ιδέα των κλίκων (βλέπε Εικόνα 3.4).

**Ορισμός 3.15 (Κλίκες).** Μια κλίκα  $C$  ενός γραφήματος  $(\Omega, \mathcal{N})$  είναι ένα σύνολο μικροστοιχείων  $\alpha, \beta \in C$ , τέτοια που για οποιαδήποτε δυο ξεχωριστά μικροστοιχεία

$$\alpha \in N_\beta \quad (\text{και σαν αποτέλεσμα της συμμετρίας, } \beta \in N_\alpha)$$

Έστω ότι  $\mathcal{C} = \mathcal{C}(\Omega, \mathcal{N})$  υποδηλώνει τη συλλογή όλων των κλίκων. Τότε το  $\mathcal{C}$  είναι φυσικά στρωματοποιημένο σύμφωνα με τα μεγέθη των κλίκων:

$$\mathcal{C} = \mathcal{C}_0 \cup \mathcal{C}_1 \cup \mathcal{C}_2 \cup \dots,$$

όπου για  $k = 0, 1, \dots$ ,

$$\mathcal{C}_k = \{C \in \mathcal{C} \mid \#C = k\}, \text{ η συλλογή όλων των κλίκων-}k$$

Προσέξτε ότι το  $\mathcal{C}_0$  περιλαμβάνει μόνο την άδεια κλίκα και το  $\mathcal{C}_1$  αποτελείται από όλες τις ελεύθερες κλίκες, και συνεπώς είναι και τα δυο ασήμαντα και ανεξάρτητα από το γειτονικό σύστημα  $\mathcal{N}$ . Το πρώτο μη-ασήμαντο στρώμα  $\mathcal{C}_2$  είναι η συλλογή όλων των διπλών κλικών. Για παράδειγμα, για τον κανονικό τομέα Καρτεσιανής  $\mathbb{Z}^2$  με το γειτονικό σύστημα προσδιορισμένο όπως στο (3.45),

$$\mathcal{C}_2 = \{(i, j), (i + 1, j)\} \mid (i, j) \in \mathbb{Z}^2\} \cup \{(i, j), (i, j + 1)\} \mid (i, j) \in \mathbb{Z}^2\}. \quad \text{Προσέξτε}$$

ότι οποιαδήποτε κλίκα  $C$  πρέπει να καλύπτεται από το  $N_\alpha$  για οποιοδήποτε  $\alpha \in C$ , και συνεπώς πρέπει να είναι τοπικό.

**Ορισμός 3.16 (Προοπτικές Κλίκας).** Μια χωρική στατιστική (ή πραγματική λειτουργική)  $V_C[u]$ , η οποία προσδίδει μια πραγματική αξία  $V_C[u]$  σε οποιοδήποτε δείγμα εικόνας  $u$ , λέγεται ότι είναι μια κλικωτή προοπτική προσαρμοσμένη σε μια κλίκα  $C$  αν για οποιαδήποτε δυο δείγματα  $u$  και  $w$ ,

$$V_C[u] = V_C[w] \text{, όποτε } u|_C = w|_C$$

Με άλλα λόγια, στην ουσία το  $V_C[u]$  είναι μόνο μια λειτουργία του  $u|_C$ .

Το αξιoσημείωτο θεώρημα των Χάμερσι και Κλίφορντ στην ισότητα των Γκιπς – Μάρκοβ δηλώνεται ως ακολούθως.

**Θεώρημα 3.17 (Ισότητα Γκιπς - Μάρκοβ).** Έστω ότι  $u$  υποδηλώνει ένα σύνολο του Γκιπς στο  $\Omega$ :

$$p(u) = \frac{1}{Z} \exp(-\beta E[u]),$$

όπως στο (3.42), και το  $\mathcal{N}$  ένα δεδομένο γειτονικό σύστημα στο  $\Omega$ . Ας υποθέσουμε ότι η ενέργεια  $E[u]$  είναι μια υπερθέση των κλικωτών προοπτικών:

$$E[u] = \sum_{C \in \mathcal{C}} V_C[u].$$

Τότε το  $u$  είναι ένα τυχαίο πεδίο Μάρκοβ στο  $(\Omega, \mathcal{N})$ . Αντιστρόφως, κάτω από μια ήπια υπόθεση θετικότητας [144], οποιοδήποτε τυχαίο πεδίο του Μάρκοβ μπορεί να σχεδιαστεί με αυτό τον τρόπο, με ένα μοναδικό σύνολο κλικωτών προοπτικών κάτω από κάποια κατάλληλη διαδικασία κανονικοποίησης.

Η απόδειξη για το πρώτο μισό είναι σχετικά ξεκάθαρη βασιζόμενη στους ορισμούς, ενώ το δεύτερο μέρος βασίζεται συχνά στη φόρμουλα του Μόππιους για αντιστροφή (συνδυαστικού) πλέγματος.

Παραπέμπουμε τον αναγνώστη στους Μπρέμαουτ [33] και Γκρίμμετ [144] για περισσότερες λεπτομέρειες στην απόδειξη.

Για να καταλήξουμε, οι τοπικές κλικωτές προοπτικές και τα σύνολα του Γκιπς συντελούν ένα βολικό μηχανήμα για το σχεδιασμό τυχαίων εικόνων. Θα επιστρέψουμε σε αυτό το θέμα στο κεφάλαιο τμηματοποίησης όταν θα επεξηγηθεί περαιτέρω η πασίγνωστη δουλειά των Γκέμαν και Γκέμαν [130].

### 3.4.5 Οπτικά Φίλτρα και Συστοιχίες Φίλτρων

οι προηγούμενες ενότητες έκαναν τις εντυπωσιακές συνδέσεις ανάμεσα σε στοχαστικά μοντέλα εικόνας και τα σύνολα του Γκιπς. Η κύρια πρόκληση τώρα είναι πώς να σχεδιάσουμε κατάλληλα τις γενικοποιημένες ενέργειες και τις συνδεδεμένες προοπτικές τους.

Αυτό που φαίνεται φυσικό και απαραίτητο είναι ότι αυτές οι ενέργειες και προοπτικές πρέπει να σχετίζονται με τα υλικά κλειδί της αντίληψης της ανθρώπινης όρασης, συμπεριλαμβανομένου, για παράδειγμα, κλίμακα, κατεύθυνση, και συχνότητα. Συνεπώς, θα τα αποκαλούμε οπτικές ενέργειες και οπτικές προοπτικές.

Από τη θεωρία άκρου του Μαρ στη σύγχρονη ανάλυση κυματιδίων, μια αξιοσημείωτη απόδοση της οπτικής αντίληψης φαίνεται να είναι η ικανότητα διαφοροποίησης. Με άλλα λόγια μια σταθερή εικόνα

$u \equiv c$  συνήθως φαίνεται μη ενδιαφέρουσα και έχει μηδενική ενέργεια. Έτσι είναι λογικό να αναμένουμε ότι για οποιαδήποτε οπτική ενέργεια  $E$ ,

$$E[c] = 0 \text{ for } c \in \mathbb{R} \text{ σαν μια σταθερή εικόνα.}$$

Αυτό εισηγείται επίσης ότι ίσως θα ήταν πιο σοφό να χτίσουμε οπτικές ενέργειες κατευθείαν στα διαστήματα χαρακτηριστικών. Έστω ότι το  $F$  είναι ένας χαρακτηριστικός τελεστής εξαγωγής ο οποίος εξάγει ένα φορέα χαρακτηριστικό ή μια λειτουργία  $Fu$  για οποιαδήποτε δεδομένη εικόνα  $u$ , και είναι υπεραπλάσιος με την έννοια ότι το  $Fc$  είναι ένας μηδενικός φορέας χαρακτηριστικού ή λειτουργία για οποιαδήποτε σταθερή εικόνα  $c$ . Τότε είναι επιθυμητό να έχουμε μια ενέργεια υπό τη μορφή του

$$E[u] = V[Fu]$$

για κάποια κατάλληλη λειτουργία (όταν το  $Fu$  είναι ένας φορέας) ή λειτουργικότητα (όταν το  $Fu$  είναι μια λειτουργία)  $V[\cdot]$  στο χαρακτηριστικό διάστημα με  $V[0] = 0$ .

**Παράδειγμα.** Ας υποθέσουμε ότι  $(\Omega, E_\Omega)$  είναι το γράφημα μικροστοιχείου που μεταφέρει όλες τις στοχευμένες εικόνες. Ας πάρουμε τη διαφοροποίηση της άκρης σαν τον τελεστή εξαγωγής

χαρακτηριστικού, δηλαδή,  $w = Fu$  είναι ο φορέας που δεικνύεται από όλα τα άκρα και προσδιορίζεται από, για οποιοδήποτε άκρο γραφήματος  $e : x \sim y$ ,

$$w_e = |u(x) - u(y)|.$$

Ας πάρουμε το  $V$  υπό τη μορφή του

$$V[w] = V_q[w] = \sum_{e \in E_\Omega} w_e^q.$$

Όταν  $q = 2$  αυτή η ενέργεια οδηγεί σε δυο σημαντικές ποσότητες του γραφήματος: τον τυχαίο πίνακα και το Λαπλασιανό γράφημα [87]. Για  $q = 1$ , η ενέργεια είναι διαχωρίσιμη τηλεοπτική ενέργεια του  $u$ .

Πιο γενικά, όπως παρακινείται από την ανάλυση όρασης, κάποιος θα μπορούσε να έχει αρκετούς τελεστές εξαγωγής χαρακτηριστικών  $(F^\alpha : \alpha \in \mathcal{A})$ , όπου  $\mathcal{A}$  είναι ένα πεπερασμένο αλφάβητο μαρκαρίσματος και τυπικά στη σειρά των δέκα [328, 329]. Έστω ότι  $u^\alpha$  υποδηλώνει το χαρακτηριστικό αποτέλεσμα  $F^\alpha u$  για κάθε  $\alpha \in \mathcal{A}$  και εικόνα  $u$ .

Στην ανάλυση εικόνας και όρασης, η εξαγωγή χαρακτηριστικών είναι συνήθως βοηθούμενη από υψιπερατά φίλτρα.

**Ορισμός 3.18 (Γραμμικά Υψιπερατά Φίλτρα).** Ένας τελεστής εξαγωγής χαρακτηριστικών  $F$  από το διάστημα εικόνας  $\mathcal{I} = \mathcal{I}(\Omega)$  σε κάποιο γραμμικό χαρακτηριστικό διάστημα  $V$  λέγεται ότι είναι γραμμικό και υψιπερατό αν

$$F(au + bv) = aFu + bFv,$$

για οποιεσδήποτε δυο σταθερές  $a$  και  $b$  και εικόνες  $u$  και  $v$ , και

$$Fc \equiv 0$$

για οποιαδήποτε σταθερή εικόνα  $c$ .

**Παράδειγμα.** Όπως και πριν ας υποθέσουμε ότι το  $(\Omega, E_\Omega)$  είναι το μεταφερόμενο γράφημα των εικόνων. Για κάθε μικροστοιχείο  $x \in \Omega$ , ορίζουμε το βαθμό του από

$$d(x) = \#\{y \in \Omega \setminus \{x\} : y \sim x\} = \#N_x,$$

δηλαδή, τον αριθμό κοντινών μικροστοιχείων με άμεση πρόσβαση άκρου στο  $\mathcal{X}$ . Ας ορίσουμε τον Λαπλασιανό τελεστή ή φίλτρο  $w = Fu$  από

$$w(x) = -d(x)u(x) + \sum_{y \in N_x} u(y). \quad (3.46)$$

Τότε είναι εύκολο να επιβεβαιώσουμε ότι το Λαπλασιανό φίλτρο είναι όντως γραμμικό και υπιπερατό. Παίζει ένα κεντρικό ρόλο στη θεωρία φασματικού γραφήματος [87]. Μια ενέργεια χτισμένη πάνω στο Λαπλασιανό φίλτρο μπορεί να δοθεί από

$$E[u] = \sum_{\Omega} |w(x)|^q$$

για μια σταθερή δύναμη  $q \geq 1$ .

Στην ανάλυση σημάτων και εικόνων, μια συστοιχία φίλτρων αναφέρεται στη συλλογή φίλτρων συχνά σχεδιασμένων για να πετύχουν ένα κοινό στόχο.

Ας υποθέσουμε ότι  $(F^\alpha : \alpha \in \mathcal{A})$  είναι μια συστοιχία γραμμικών υπιπερατών φίλτρων ικανών να εξάγουν όλα τα κύρια οπτικά χαρακτηριστικά για τα οποία νοιάζεται ένα φυσιολογικό σύστημα ανθρώπινης όρασης. Αυτό σημαίνει ότι θα έπρεπε να μπορεί να επιλύει χαρακτηριστικά κλειδί όπως οι τοπικές χωρικές συχνότητες, προσανατολισμοί, κλπ. Τότε θα ήταν λογικό να υποθέσουμε ότι ένα πιστό στοχαστικό μοντέλο θα εξαρτόταν μόνο από τις αντιδράσεις του φίλτρου μιας τέτοιας συστοιχίας φίλτρων. Αλλά το ερώτημα είναι πώς να χαρακτηρίσουμε μια τέτοια εξάρτηση.

Μια προσέγγιση, της οποίας η γένεση και αυστηρή παραγωγή θα δοθεί στην επόμενη υποενότητα, υποθέτει επίσημα ότι όλες οι εξαγωγές φίλτρων  $(u^\alpha : \alpha \in \mathcal{A})$  είναι κάπως ανεξάρτητες, και συνεπώς η ξεχωριστές τους ενέργειες  $E_\alpha[u^\alpha]$  συμπεριφέρονται σαν τις ανεξάρτητες εκτεταμένες μεταβλητές  $E$  (εσωτερική ενέργεια),  $V$  (ένταση), και  $n$  (γραμμοριακοί αριθμοί) σε ένα τυπικό GCE στη στατιστική μηχανική (βλέπε Ενότητα 2.3). Με άλλα λόγια, το μοντέλο εικόνων Γκιπς πέρνει την ακόλουθη μορφή:

$$p(u) = \frac{1}{Z} \exp \left( - \sum_{\alpha \in \mathcal{A}} \beta_\alpha E_\alpha[u^\alpha] \right), \quad (3.47)$$

όπου

$\beta_\alpha$  είναι συνδεδεμένες προοπτικές. Για την ακρίβεια, ακόμη πιο γενικά, για κάθε τάξη

χαρακτηριστικού  $\alpha \in \mathcal{A}$ , κάποιος θα μπορούσε να επιτρέψει τόσο το  $E_\alpha$  όσο και το  $\beta_\alpha$  να είναι φορείς:

$$E_\alpha = [E_\alpha^1, E_\alpha^2, \dots, E_\alpha^{m_\alpha}] \text{ and } \beta_\alpha = [\beta_\alpha^1, \beta_\alpha^2, \dots, \beta_\alpha^{m_\alpha}].$$

Τότε το

μοντέλο εικόνας Γκιπς δίνεται από

$$p(u) = \frac{1}{Z} \exp \left( - \sum_{\alpha \in \mathcal{A}} \beta_\alpha \cdot E_\alpha[u^\alpha] \right). \quad (3.48)$$

Αυτή

είναι η πιο γενική μορφή των μοντέλων εικόνων του Γκιπς στη λογοτεχνία.

### 3.4.6 Μάθηση Μοτίβων Εικόνας Βασισμένη στην Εντροπία

Σε αυτή την ενότητα θα κάνουμε μια σύντομη εισαγωγή στην προσέγγιση των Ζου, Μάμφορτ και Γου σχετικά με την εκμάθηση μοντέλων εικόνων του Γκιπς βασισμένα στην αρχή της μέγιστης εντροπίας [328, 329].

Ακολουθώντας το γενικό μοντέλο του Γκιπς (3.48) στην προηγούμενη υποενότητα, οδηγούμαστε στα ακόλουθα τρία ερωτήματα:

1. Τι είναι μια καλή συστοιχία φίλτρων  $(F^\alpha : \alpha \in \mathcal{A})$  που παγιδεύει οπτικά χαρακτηριστικά αρκετά καλά;

2. Ποιες είναι οι οπτικές ενέργειες (επεκταμένες μεταβλητές)  $E_\alpha$  για κάθε κανάλι χαρακτηριστικής εξαγωγής  $u^\alpha = \mathcal{F}^\alpha u$ ;

3. Ποιες είναι οι κατάλληλες οπτικές προοπτικές (ή δυαδικές εντατικές μεταβλητές)  $\beta_\alpha$  που είναι συνδεδεμένες με τις ενέργειες;

Ας ανακαλέσουμε όπως και στην Ενότητα 2.3 ότι η ποσότητα  $X = X[u|\Omega]$  λέγεται ότι είναι εκτατική αν για οποιοσδήποτε δυο διαχωρισμένους τομείς εικόνων  $\Omega_1$  και  $\Omega_2$ , και μια εικόνα  $u$  σε «ισορροπία»,

$$X[u|\Omega_1 \cup \Omega_2] = X[u|\Omega_1] + X[u|\Omega_2].$$

Παρομοίως, το  $X$  αποκαλείται μια εντατική μεταβλητή αν αντί για αυτό.

$$X[u|\Omega_1 \cup \Omega_2] = X[u|\Omega_1] = X[u|\Omega_2].$$

Ακολουθώντας τη στατιστική μηχανική, θα έπρεπε και τα δυο να γίνουν κατανοητά με την ασυμπτωτική έννοια όταν τα μεγέθη των τομέων είναι αρκετά μεγάλα.

Από την άλλη, σε αντίθεση με τη στατιστική μηχανική όπου τα ισοζύγια (π.χ., θερμικά, μηχανικά, ή χημικά) είναι φυσικά προσδιορισμένα μέσω τις φυσικής επαφής (π.χ., μέσω θερμικής επαφής, ένα μηχανικό έμβολο ή μια διαπερατή μεμβράνη), ο ορισμός του ισοζυγίου των δυο τμημάτων μιας εικόνας ή δυο διαφορετικών εικόνων είναι πού λιγότερο εμφανές. Αυτό που φαίνεται φυσικό για την ανάλυση όρασης και εικόνων είναι ότι η ιδέα του ισοζυγίου πρέπει να είναι με την έννοια τις οπτικής διάκρισης. Διαισθητικά μιλώντας, δυο εικόνες λέγεται ότι είναι σε (οπτικό) ισοζύγιο αν η μια δεν μπορεί να διακριθεί από την άλλη όταν βρίσκονται δίπλα δίπλα.

Στο πιο ιδανικό σενάριο, μια τάξη φυσικών εικόνων χαρακτηρίζεται καλύτερα από μια διαδικασία μάθησης που θα μπορούσε να μάθει καλύτερα όλα τα τρία στοιχεία της μοντελοποίησης του Γκιπς: τα οπτικά φίλτρα  $(F^\alpha : \alpha \in \mathcal{A})$ , οι εκτεταμένες μεταβλητές «ενέργειας»

$(E_\alpha : \alpha \in \mathcal{A})$ , και οι εντατικές οπτικές προοπτικές  $(\beta_\alpha : \alpha \in \mathcal{A})$ . Μια τέτοια μαθησιακή διαδικασία, όμως, είναι υπερβολικά ακριβή.

Στο έργο των Ζου και Μάμφορτ [328] και Ζου, Γου και Μάμφορτ [329], προτείνεται ένα μαθησιακό μοντέλο για τις οπτικές προοπτικές, δεδομένου ότι τα δυο λεξικά δίνονται ή επιλέγονται εκ των προτέρων για τα οπτικά φίλτρα και τις εκτεταμένες μεταβλητές ενέργειας. Είναι όμορφα ριζωμένο στην διαφοροποιημένη διατύπωση του Γκιπς της στατιστικής μηχανικής – η αρχή της μέγιστης εντροπίας, όπως εισήχθηκε στην προηγούμενη Ενότητα 2.3.

Η συστοιχία φίλτρων που χρησιμοποιείται αποτελείται από γραμμικά φίλτρα πολλαπλής κλίμακας Γκαμπορ. Έτσι για παράδειγμα ένας δείκτης χαρακτηριστικών  $\alpha$  είναι συνδεδεμένος με μια τριπλή κλίμακα  $\sigma > 0$ , προσανατολισμό  $\theta \in (-\pi, \pi]$ , και χωρική συχνότητα  $k > 0$ :

$$\alpha = (\sigma, \theta, k) \in \mathbb{R}^3$$

και το  $\mathcal{A}$  είναι ένα ποσοστοποιημένο πεπερασμένο υποσύνολο στο  $\mathbb{R}^3$ . Τότε για κάθε

$\alpha \in \mathcal{A}$ ,  $u^\alpha = F^\alpha u$  είναι μια κατανομή χαρακτηριστικών στον τομέα κλιμάκωσης

$\Omega_\alpha = \Omega_\sigma$ , ο οποίος δίνεται από, για παράδειγμα,

$$\Omega_2 = \{(2i, 2j) \in \mathbb{Z}^2 : 0 \leq i \leq (I-1)/2 \text{ and } 0 \leq j \leq (J-1)/2\},$$

$$\Omega_4 = \{(4i, 4j) \in \mathbb{Z}^2 : 0 \leq i \leq (I-1)/4 \text{ and } 0 \leq j \leq (J-1)/4\}, \dots$$

υποθέτοντας ότι ο αρχικός τομέας εικόνας είναι  $\Omega = (0 : I-1) \times (0 : J-1)$ .

Προσέξτε ότι εδώ οι δείκτες κλίμακας 2 και 4 είναι οι κλίμακες σμίκρυνσης. Στη πράξη, συχνά πολλά επίπεδα κλίμακας είναι αρκετά για γενική ανάλυση εικόνας. Για την ώρα, θα αγνοήσουμε τα αποτελέσματα των ορίων χρησιμοποιώντας αντανακλαστικές ή περιοδικές επεκτάσεις. Επιπρόσθετα,

θα αποκαλέσουμε το  $\Omega_\alpha$ , τον φορέα του χαρακτηριστικού τύπου  $\alpha$ .

Οι μεταβλητές εκτεταμένης ενέργειας στους Ζου, Μάμφορτ, και Γου η μοντελοποίηση εικόνων του Γκιπ είναι τα ιστογράμματα των κατανομών χαρακτηριστικών.

Πρώτα, η εμβέλεια κάθε κατανομής χαρακτηριστικού  $u^\alpha$  είναι εξίσου κβαντισμένη με συγκεκριμένα πεπερασμένα επίπεδα (υποδηλωμένα από το σύμβολο τμηματοποίησης  $\pi$ ):

$$\pi : \dots < b < a < c < \dots ,$$

όπου το βήμα κβαντοποίησης ή το μέγεθος της αποθήκης  $h = a - b = c - a = \dots$  μπορεί να εξαρτάται από τον τύπο χαρακτηριστικών  $\alpha$ . Για κάθε δείκτη αποθήκης  $a$ , ορίζουμε τη ζώνη επιπέδου  $\Omega_\alpha^a$  από

$$\Omega_\alpha^a = \{x \in \Omega_\alpha : a - h/2 \leq u^\alpha(x) < a + h/2\}.$$

Τότε ο (ακανόνιστος) φορέας ιστογράμματος  $H_\alpha$  για τον τύπο χαρακτηριστικών  $\alpha$  υπό το σχήμα κβαντοποίησης  $\pi$  προσδιορίζεται ως

$$H_\alpha = (\dots, H_\alpha^b, H_\alpha^a, H_\alpha^c, \dots), \text{ with } H_\alpha^a = \#\Omega_\alpha^a. \quad \text{Προσέξτε ότι για}$$

κάθε τύπο χαρακτηριστικού  $\alpha$ , το  $H_\alpha = H_\alpha[u]$  είναι πράγματι μια λειτουργία (μη-γραμμικού) φορέα της αρχικής εικόνας  $u$ . Στο μοντέλο των Ζου, Μάμφορτ και Γου [328. 329], αυτά

λαμβάνονται ως οι ενέργειες  $E_\alpha$  για το γενικό μοντέλο συνόλου(3.48). Είναι κρίσιμο να

επιβεβαιώσουμε ότι κάθε  $H_\alpha$  επεκτείνεται, όταν ο τύπος φίλτρου  $\alpha$  και το σχήμα κβαντοποίησης δίνονται.

Τέλος, οδηγούμαστε στον κατάλληλο προσδιορισμό των συσχετισμένων προοπτικών. Για το λόγο ο οποίος θα ξεκαθαρίσει αμέσως στη συνέχεια, αυτές οι προοπτικές θα υποδηλωθούν από

$\lambda_\alpha(a)$  αντί για τη γενική εγγραφή  $\beta_\alpha^a$  στην προηγούμενη υποενότητα και μαθαίνονται από τη σύνθεση της μέγιστης εντροπίας του Γκιπς.

Πρώτα, προσέξτε ότι τα ιστογράμματα  $H_\alpha$  είναι τυχαίοι φορείς εφόσον οι εικόνες θεωρούνται ως τυχαία πεδία. Ας υποθέσουμε ότι έχουμε μια επαρκώς μεγάλη συστοιχία  $C$  μιας τάξης – στόχου φυσικών εικόνων. Τότε η στατιστική των ιστογραμμάτων μπορεί να εξασφαλιστεί από εμπειρικά σύνολα ολοκλήρωσης, για παράδειγμα, το μέσο

$$\langle H_\alpha \rangle_{(\mathcal{I}, p)} \simeq \langle H_\alpha[u] \rangle_{u \in C} = h_\alpha, \quad \alpha \in \mathcal{A}. \quad (3.49)$$

Αν αντι

για αυτού, αυτά επιβληθούν ως ακριβείς περιορισμοί στην τυχαία κατανομή πεδίων  $p(u)$  για να μοντελοποιηθούν, σαν ανάλογα στην σύνθεση GCE (2.32) στη στατιστική μηχανική, τότε η αρχή της μέγιστης εντροπίας του Γκιπς, οδηγεί φυσικά στο μοντέλο πεδίου του Γκιπς



$$p(u) = p_{\Lambda}(u) = \frac{1}{Z} \exp \left( - \sum_{\alpha \in \mathcal{A}} \lambda_{\alpha} \cdot H_{\alpha}[u] \right),$$

όπου το  $\Lambda = (\lambda_{\alpha} : \alpha \in \mathcal{A})$  μαζεύει όλους τους πολλαπλασιαστές Λαγκραντζ για την περιορισμένη μεγιστοποίηση της εντροπίας του Γκιπς (2.31), οι οποίοι αποκαλούνται οπτικές προοπτικές.

Για να καθορίσουμε ή να μάθουμε αυτές τις οπτικές προοπτικές, κάποιος προσέχει ότι η λειτουργία τμηματοποίησης εξαρτάται από τις οπτικές προοπτικές

$$Z = Z_{\Lambda} = \sum_{u \in \mathcal{I}} \exp \left( - \sum_{\alpha \in \mathcal{A}} \lambda_{\alpha} \cdot H_{\alpha}[u] \right),$$

όπου έχει υποτεθεί ότι οι γκριζες κλίμακες εικόνων κβαντοποιούνται επίσης (π.χ., κβαντοποίηση 8 ή 16 bits), και συνεπώς το  $\mathcal{I} = \mathcal{I}(\Omega)$  αποτελείται από πεπερασμένα πολλές εικόνες. Επιπλέον, έχουμε όπως στη στατιστική μηχανική (Ενότητα 1.3.3)

$$\frac{-\partial \ln Z_{\Lambda}}{\partial \lambda_{\alpha}} = \langle H_{\alpha} \rangle_{(\mathcal{I}, p_{\Lambda})} = h_{\alpha}, \quad \alpha \in \mathcal{A}. \quad (3.50)$$

Αυτές είναι οι εξισώσεις για τις άγνωστες οπτικές προοπτικές  $\Lambda = (\lambda_{\alpha} : \alpha \in \mathcal{A})$  και μπορούν να λυθούν στατιστικά χρησιμοποιώντας τους δειγματιστές του Γκιπς.

Γράφουμε  $\lambda_{\alpha}(a) = \tilde{\lambda}_{\alpha}^a$  για οποιαδήποτε επίπεδο κβαντοποίησης  $a$  της τάξης  $\lambda_{\alpha}(y)$ , και το επεκτείνουν σε μια σημειακή σταθερή λειτουργία  $\lambda_{\alpha}(y)$  που συμβαδίζει με το σχήμα κβαντοποίησης  $\pi : \dots < b < a < c < \dots : \lambda_{\alpha}(y) = \tilde{\lambda}_{\alpha}^a$  για οποιοδήποτε  $y \in [(b+a)/2, (a+c)/2)$ . Τότε

$$\lambda_{\alpha} \cdot H_{\alpha} = \sum_{\dots < b < a < c < \dots} \lambda_{\alpha}(a) \times \#\Omega_{\alpha}^a = \sum_{y \in \Omega_{\alpha}} \lambda_{\alpha}(u^{\alpha}(y)).$$

Έτσι για συνεχόμενους τομείς, με ένα πιθανό σταθερό πολλαπλασιαστή, γίνεται

$$\lambda_{\alpha} \cdot H_{\alpha} \rightarrow \int_{\Omega_{\alpha}} \lambda_{\alpha}(u^{\alpha}(y)) dy,$$

το οποίο οδηγεί στη γνώριμη μορφή των «ελεύθερων ενεργειών» που χρησιμοποιούνται από την κοινότητα μεταβλητής ΛΣΕ επεξεργασίας εικόνας:

$$\sum_{\alpha \in \mathcal{A}} \lambda_{\alpha} \cdot H_{\alpha} \rightarrow \sum_{\alpha \in \mathcal{A}} \int_{\Omega_{\alpha}} \lambda_{\alpha}(u^{\alpha}(y)) dy.$$

Για παράδειγμα, επιλέγοντας εκ των προτέρων την κλίση υψιπερατού φίλτρου

$$u^{\alpha}(y) = |\nabla u(y)|, \quad y \in \Omega = \Omega_{\alpha},$$

και η οπτική προοπτική  $\lambda_{\alpha}(a) \equiv a$  οδηγεί επίσης στο μοντέλο εικόνας BV των Ρούτιν, Όσερ και Φατέμι [258], και Ρούτιν και Όσερ [257] που παρουσιάστηκαν στο προηγούμενο κεφάλαιο.

### 3.5 Αναπαράσταση Επιπέδου-Ομάδας

Μια εικόνα σαν λειτουργία μπορεί να γίνει κατανοητή ως μια συλλογή ισόφωτων, ή ισοδύναμα, επιπέδων ομάδων. Αυτή η άποψη οδηγεί στην αναπαράσταση επιπέδου-ομάδας εικόνων και είναι στενά συνδεδεμένη στην πασίγνωστη υπολογιστική τεχνολογία επιπέδου-ομάδας των Όσερ και Σέθιαν [241], Όσερ και Φέτκιου [243], Όσερ και Παραγιός [239], και Σέθιαν [269]. Πρέπει να δώσουμε έμφαση στο γεγονός ότι η αναπαράσταση επιπέδου-ομάδας μπορεί επίσης να θεωρηθεί ως η συνέχιση του ορίου της μορφολογικής προσέγγισης στην κλασσική επεξεργασία διακριτικής εικόνας, όπως αναφέρθηκε σε συντομία στο Κεφάλαιο 1 (δείτε επίσης, π.χ., [48, 147, 267]).

#### 3.5.1 Κλασσικά Επίπεδα Ομάδων

Ας θεωρήσουμε μια εικόνα  $u = u(x)$  γκριζας κλιμακας ως μια λειτουργία σε ένα περιορισμένο  $x \in \Omega$ ,

δυσδιάστατο ανοιχτό τομέα ο οποίος είναι τυπικά ένας τετράγωνος τομέας στις πλείστες ψηφιακές εφαρμογές.

Για κάθε πραγματική αξία  $\lambda$ , ορίζουμε το επίπεδο ομάδας  $\lambda$  του  $u$  να είναι

$$\gamma_{\lambda} = \{x \in \Omega : u(x) = \lambda\}. \quad (3.51)$$

Τότε η κλασσική αναπαράσταση επιπέδου – ομάδας του  $u$  είναι η μονοπαραμετρική οικογένεια των επιπέδων ομάδων

$$\Gamma_u = \{\gamma_{\lambda} : \lambda \in \mathbb{R}\}. \quad (3.52)$$

Προσέξτε ότι το  $\Gamma_u$  είναι μια τμηματοποίηση του τομέα εικόνας, εννοώντας ότι

$$\Omega = \cup_{\lambda \in \mathbb{R}} \gamma_{\lambda}, \quad \gamma_{\lambda} \cap \gamma_{\mu} = \emptyset, \quad \lambda \neq \mu.$$

Ας υποθέσουμε πρώτα ότι το  $u$  είναι αρκετά ομαλό (μετά από μια απειροελάχιστη διάχυση θερμότητας, ας πούμε). Τότε η διαφοροποιημένη τοπολογία είναι βοηθητική στην κατανόηση της γενικής συμπεριφοράς της αναπαράστασης του επιπέδου-ομάδας [218].

Μια αξία  $\lambda$  λέγεται ότι είναι κανονική αν η κλίση του  $u$  δεν εξαφανίζεται ποτέ πουθενά κατά μήκος του  $\gamma_\lambda$ . Αν ισχύει αυτό, το επίπεδο ομάδας  $\gamma_\lambda$  πρέπει να είναι ένα μονοδιάστατο υποσυγκρότημα του  $\Omega$ , που σημαίνει ότι τοπικά φαίνεται σαν (μετά από μια διαφοροποιημένη αποσύνθεση) το διάστημα  $(-1, 1) \times \{0\}$  ενσωματωμένο στο τετράγωνο  $Q = (-1, 1) \times (-1, 1)$ , χωρίς διακλάδωση ή απόλιξη. Για ευκολία, ένα τέτοιο μονοδιάστατο υποσυγκρότημα αποκαλείται συχνά απλά μια κανονική καμπύλη, παρόλο που μπορεί να αποτελείται από πολλές συνδεδεμένες συνιστώσες.

Σε διαφορετική περίπτωση, μια αξία  $\lambda$  λέγεται ότι είναι ιδιάζουσα. Οποιοδήποτε μικροστοιχείο  $x_0 \in \gamma_\lambda$  με  $\nabla u(x_0) = 0$  θεωρείται συνεπώς κρίσιμο. Ένα κρίσιμο σημείο είναι όπου βρίσκεται ένα τοπικό ελάχιστο ή μέγιστο (κοιλιάδα ή κορυφή). Θα μπορούσε να είναι επίσης ένα διακριτικό σημείο όπου προς τη μια κατεύθυνση η εικόνα γίνεται πιο σκούρα ενώ κατα μήκος της κάθετης της γίνεται πιο φωτεινή. Αυτές είναι μόνο τρεις γενικές περιπτώσεις σε δυο διαστάσεις. Επιπλέον, ένα γενικό κρίσιμο σημείο απομονώνεται πάντοτε.

Από τη διαφοροποιητική τοπολογία [218], η ομάδα κανονικών αξιών μιας ομαλής εικόνας  $u$  είναι ανοιχτό και πυκνό στο  $\mathbb{R}$ . Δηλώνει ότι για οποιαδήποτε αξία  $\lambda$  κάποιος επιλέγει τυχαία, είναι σχεδόν σίγουρο (με την έννοια Λεπέσκ) ότι το συνδεδεμένο επίπεδο ομάδας  $\gamma_\lambda$  είναι μια κανονική καμπύλη.

### 3.5.2 Σωρευτικά Επίπεδα Ομάδων

Στην πραγματικότητα, οι εικόνες σπάνια είναι ομαλές λειτουργίες. Αντιθέτως υποθέτονται ότι είναι σε συγκεκριμένα λειτουργικά διαστήματα όπως  $L^1(\Omega)$  ή  $L^2(\Omega)$ , κλπ. Τότε οι σημειακές αξίες δεν είναι τόσο σημαντικές εφόσον είναι πάντα προσαρμόσιμες πάνω από μια ομάδα μηδενικής μέτρησης (Λεπέσκ). Έτσι το ξεχωριστό επίπεδο ομάδων ίσως έχει λιγότερο νόημα.

Το κλασσικό σημειακό επίπεδο ομάδων μπορεί να προσαρμοστεί στο σωρευτικό επίπεδο ομάδων  $F_\lambda$  προσδιορίζεται από το

$$F_\lambda = F_\lambda(u) = \{x \in \Omega : u(x) \leq \lambda\}. \quad (3.53)$$

Κάθε σωρευτικό επίπεδο ομάδων προσδιορίζεται σωστά στη μετρική-θεωρητική έννοια εφόσον αν το  $u(x) = v(x)$ , σχεδόν σίγουρα  $x \in \Omega$ , η συμμετρική διαφορά

$$F_\lambda(u) \ominus F_\lambda(v) = (F_\lambda(u) - F_\lambda(v)) \cup (F_\lambda(v) - F_\lambda(u))$$

έχει επίσης μηδενική μέτρηση. Η αναπαράστατη του σωρευτικού επιπέδου-ομάδας μιας δεδομένης εικόνας  $u$  είναι η οικογένεια μιας παραμέτρου

$$\mathcal{F}_u = \{F_\lambda : \lambda \in \mathbb{R}\}.$$

Προσέξτε ότι οι δυο ιδέες του επιπέδου ομάδων σχετίζονται από το

$$F_\lambda = \cup_{\mu \leq \lambda} \gamma_\mu$$

Αν η εικόνα  $u$  είναι μια συνεχόμενη λειτουργία στο  $\Omega$  κάθε  $F_\lambda$  είναι μια (σχετικά) κλειστή ομάδα και το τοπολογικό της όριο  $\partial F_\lambda = \gamma_\lambda$ . Για μια γενική μετρήσιμη εικόνα Λεπέσκ, κάθε ένα από τα επίπεδα ομάδων της πρέπει να είναι επίσης μετρήσιμο.

Η κανονικότητα ενός επιπέδου ομάδας  $\gamma_\lambda$  ή  $F_\lambda$  μπορεί να μετρηθεί με πολλούς τρόπους. Η χαμηλότερη τάξη είναι η μέτρηση μήκους μήκος  $(\gamma_\lambda)$ , η οποία είναι καλά προσδιορισμένη αν το  $\gamma_\lambda$  είναι ομαλό ή Λιπσχιτζ.

Πιο γενικά, η ιδέα του μήκους επεκτείνεται στην περίμετρο μέσω σωρευτικών επιπέδων ομάδων

$$\text{Per}(F_\lambda | \Omega) = \int_{\Omega} |D\chi_{F_\lambda}|, \quad (3.54)$$

δηλαδή, η τηλεόραση του δείκτη λειτουργείας. Από την φόρμουλα συν-περιοχής στην Ενότητα 2.2, κάποιος έχει

$$\int_{\Omega} |Du| = \int_{-\infty}^{\infty} \text{Per}(F_\lambda | \Omega) d\lambda. \quad (3.55)$$

Όταν το  $u$  είναι ομαλό, αυτό γίνεται απλά

$$\int_{\Omega} |\nabla u| dx = \int_{-\infty}^{\infty} \text{length}(\gamma_\lambda) d\lambda.$$

Για την αναπαράσταση του επιπέδου-ομάδας, η όμορφη ταυτότητα (3.55) λόγω των Ντε Τζόρτζι [134] και Φλέμιγκ και Ρίσελ [125] είναι ανάλογη της διάσημης ταυτότητας του Παρσέβαλ στην ανάλυση του Φουριέ [193],

$$\int_{\mathbb{R}^2} |u(x)|^2 dx = \int_{\mathbb{R}^2} |\hat{u}(\omega)|^2 d\omega,$$

όπου ο δυσδιάστατος μετασχηματισμός του Φουριέ προσδιορίζεται ως

$$\hat{u}(\omega) = \int_{\mathbb{R}^2} u(x) e^{-i2\pi\omega \cdot x} dx, \quad \omega \cdot x = \omega_1 x_1 + \omega_2 x_2.$$

Η αναλογία γίνεται πιο σαφής με την αντιστοιχία,

$$\lambda \longleftrightarrow \omega,$$

$$F_\lambda \text{ or } \gamma_\lambda \longleftrightarrow \hat{u}(\omega),$$

$$\text{Per}(\cdot) \text{ or } \text{length}(\cdot) \longleftrightarrow |\cdot|^2.$$

Έτσι η αναπαράσταση του επιπέδου-ομάδας διατηρεί την τηλεοπτική έννοια, ακριβώς όπως κάνει η

αναπαράσταση του Φουριέ στην ενέργεια  $L^2$ . Αλλά σε αντίθεση με το μετασχηματισμό του Φουριέ αναπαράσταση επιπέδου-ομάδας είναι μη γραμμική.

### 3.5.3 Σύνθεση Επιπέδου-Ομάδας

Υπό την αναπαράσταση επιπέδου-ομάδας, οποιαδήποτε δεδομένη εικόνα  $u$  μετασχηματίζεται σε μια μονοπαραμετρική οικογένεια (σωρευτικού) επιπέδου ομάδων

$\mathcal{F}_u = \{F_\lambda(u) : \lambda \in \mathbb{R}\}$ . Αυτό είναι το βήμα ανάλυσης. Τώρα αναλύουμε το βήμα σύνθεσης, δηλαδή, πώς να συνθέσουμε μια εικόνα από τις πληροφορίες των επιπέδων ομάδων.

Αντιθέτως, δεδομένης οποιασδήποτε τέτοιας μονοπαραμετρικής οικογένειας των μετρήσιμων ομάδων

του Λεπέσκ  $\mathcal{F} = \{F_\lambda : \lambda \in \mathbb{R}\}$ , με την προϋπόθεση ότι πληρεί τις δυο ακόλουθες συνθήκες συμβατότητας:

1. για οποιαδήποτε  $\lambda \leq \mu$ ,  $F_\lambda \subseteq F_\mu$  (μονοτονικότητα) και (δεξιά συνέχιση)

$$F_\lambda = \bigcap_{\mu > \lambda} F_\mu = \lim_{\mu \rightarrow \lambda^+} F_\mu,$$

2.  $F_{-\infty} := \lim_{\lambda \rightarrow -\infty} F_\lambda$  είναι ένα κενό σύνολο και

$$F_\infty := \lim_{\lambda \rightarrow \infty} F_\lambda = \Omega,$$

πρέπει να υπάρχει μια μοναδική μετρίσιμη εικόνα Λεπέσκ  $u$  στο  $\Omega$  τέτοιο που  $\mathcal{F}_u = \mathcal{F}$ . Η μοναδικότητα είναι βέβαια με την μετρική-θεωρητική έννοια. Ας αναλύσουμε αυτό το θεώρημα με περισσότερες λεπτομέρειες.

Η σύνθεση του  $\mathbf{u}$  μπορεί να πραγματοποιηθεί από, για οποιοδήποτε δεδομένο μικροστοιχείο  $x \in \Omega$ ,

$$u(x) = \inf\{\mu : x \in F_\mu\}. \quad (3.56)$$

Η δεύτερη συνθήκη συμβατότητας εγγυάται ότι το  $u(x)$  είναι καλά προσδιορισμένο παντού στο  $\Omega$  και πάντα πεπερασμένο.

Τώρα δείχνουμε ότι για την εικόνα  $\mathbf{u}$  συντεθιμένη από το (3.56),

$$F_\lambda(u) = F_\lambda \quad \text{for any } \lambda \in \mathbb{R}.$$

Πρώτα, το  $x \in F_\lambda(u)$  δηλώνει ότι  $u(x) \leq \lambda$ . Από τον ορισμό (3.56),  $x \in F_\mu$  για οποιοδήποτε  $\mu > \lambda$ . Τότε από την πρώτη συνθήκη συμβατότητας,  $x \in F_\lambda$  δηλώνοντας ότι  $F_\lambda(u) \subseteq F_\lambda$ . Αντιθέτως, αν  $x \in F_\lambda$ , κατά ορισμό,  $u(x) \leq \lambda$ . Ως αποτέλεσμα  $x \in F_\lambda(u)$  και  $F_\lambda \subseteq F_\lambda(u)$ . Αυτό αποδεικνύει ότι  $F_\lambda(u) \equiv F_\lambda$  για οποιοδήποτε  $\lambda \in \mathbb{R}$ . Συγκεκριμένα, η συντεθιμένη εικόνα είναι πράγματι μετρίσιμη κατά Λεπέςκ.

Για να δείξουμε τη μοναδικότητα, ας υποθέσουμε ότι δυο εικόνες  $\mathbf{u}$  και  $\mathbf{v}$  ικανοποιούν

$$F_\lambda(u) \equiv F_\lambda \equiv F_\lambda(v) \quad \text{for any } \lambda \in \mathbb{R}.$$

Τότε για οποιοδήποτε μικροστοιχείο  $x \in \Omega$ ,  $u(x) \leq \lambda$  μόνο και μόνο αν  $v(x) \leq \lambda$ . Έτσι αν υπάρχει κάποιο μικροστοιχείο  $x \in \Omega$  τέτοιο που  $u(x) < v(x)$ , πέρνοντας το  $\lambda = (u(x) + v(x))/2$  θα οδηγήσει σε άμεση αντιπαράθεση.

### 3.5.4 Ένα Παράδειγμα: Επίπεδα Ομάδων Τμηματικά Σταθερών Εικόνων

Ως παράδειγμα, ας θεωρήσουμε μια τμηματικά σταθερή εικόνα  $\mathbf{u}$  που δίνεται από το

$$u(x) = \sum_{n=1}^N c_n \chi_{\Omega_n}(x), \quad \text{with } c_1 \leq c_2 \leq \dots \leq c_N,$$

όπου το  $\Omega = \cup_{n=1}^N \Omega_n$  είναι ένας διαχωρισμός. Για μια τέτοια τάξη εικόνων, η σωρευτική αναπαράσταση επιπέδου-συνόλου δίνεται αποκλειστικά από το

$$F_\lambda = \cup_{m \leq n} \Omega_m \text{ for } c_n \leq \lambda < c_{n+1},$$

με την υπόθεση ότι  $c_0 = -\infty$  και  $c_{N+1} = +\infty$ .

Ορίζουμε το

$$s(\lambda) = \text{Per}(F_\lambda | \Omega)$$

να είναι η περίμετρος κάθε επιπέδου-συνόλου του  $u$ . Έστω ότι

$$s_n = \text{Per}(\cup_{m \leq n} \Omega_m | \Omega) \text{ for } 1 \leq n \leq N.$$

Προσέξτε ότι  $s_N = \text{Per}(\Omega | \Omega) = 0$ . Τότε  $s(\lambda)$  ως μια τμηματική σταθερή λειτουργία δίνεται από

$$s(\lambda) = \sum_{n=1}^N s_n \chi_{[c_n, c_{n+1})}(\lambda).$$

Έτσι, από τη φόρμουλα συν-περιοχής (3.55) (και  $s_N = 0$ ),

$$\int_{\Omega} |Du| = \int_{-\infty}^{\infty} s(\lambda) d\lambda = \sum_{n=1}^{N-1} s_n (c_{n+1} - c_n).$$

Κάποιος θα μπορούσε να δείξει εύκολα ότι η τελευταία εξίσωση είναι ίση με το ζύγισμα του μήκους κάθε αλματικού τμήματος από το μέγεθος του άλματός του στην αρχική εικόνα  $u$ .

Από την άποψη της ρεαλιστικής κωδικοποίησης, για τέτοιες τμηματικές σταθερές εικόνες, είναι

πλεονάζον να προσπαθούμε να κωδικοποιήσουμε όλα τα επίπεδα συνόλων  $F_\lambda$  εφόσον τα πλείστα από αυτά είναι πανομοιότυπα. Αν η μονάδα τηλεόρασης γίνει αποδεκτή ως μια πιστή οπτική μέτρηση, η τελευταία ταυτότητα εισηγείται την κωδικοποίηση μόνο των διευθύνσεων όλων των αλματικών

τμημάτων (υποθέτωντας ότι κάθε  $\Omega_n$  είναι Λίπσχιτζ) και τα συνδεδεμένα τους σεσημασμένα άλματα. Θα μπορούσε τότε να αποδειχθεί ότι σε μια αλλαγή ομοιόμορφης γκρίζας κλίμακας, η αρχική τμηματική σταθερή εικόνα  $u$  μπορεί να αποκωδικοποιηθεί ή να συντεθεί τέλεια.

### 3.5.5 Κανονικότητα Υψηλής Τάξης Επιπέδων Συνόλων

Σε ορισμένες εφαρμογές όπως η μη συρρίκνωση [234] και η επαναζωγράφιση εικόνων [61, 116], το μήκος ή η περίμετρος ενέργειας για επίπεδα συνόλων είναι ανεπαρκής για οπτικά πιστά αποτελέσματα. Τότε, όπως και στην κλασική θεωρία παρεμβολής, κάποιος πρέπει να λάβει υπόψη του γεωμετρικές κανονικότητες υψηλής τάξης.

Σε αυτή την υποενότητα το  $u$  θα θεωρηθεί λείο (τουλάχιστον  $C^2$ , ας πούμε). Έστω ότι το  $\lambda$  είναι μια κανονική αξία και το  $\gamma_\lambda$  το συνδεδεμένο επίπεδο συνόλου, το οποίο είναι μια λεία κανονική καμπύλη στο  $\Omega$ . Η φυσική σεσημασμένη καμπύλωση του  $\gamma_\lambda$  σε οποιοδήποτε από τα μικροστοιχεία του  $x$  δίνεται από

$$\kappa = \kappa(x|\gamma_\lambda) = \nabla \cdot \left[ \frac{\nabla u}{|\nabla u|} \right].$$

Κάποιος θα μπορούσε να επιβεβαιώσει εύκολα ότι το  $\kappa$  είναι αρνητικό κοντά στην κορυφή (ή τοπικό μέγιστο) ενώ είναι θετικό κοντά στο ελάχιστο, εφόσον το κανονικοποιημένο πεδίο κλίσης

$\nabla u / |\nabla u|$  είναι συσταλτικό και επεκτατικό, αντίστοιχα.

Μελετώντας το γεωμετρικό σχήμα ενός μη-στρεπικού λεπτού ελαστικού ράβδου, ο Έουλερ το 1744 [196] εισήγαγε για πρώτη φορά την ενέργεια καμπύλης υδροστατικού σχήματος για μια δυσδιάστατη καμπύλη  $\gamma$ :

$$e[\gamma] = \int_{\gamma} (\alpha + \beta \kappa^2) ds, \quad (3.57)$$

όπου

$\alpha$  και  $\beta$  είναι κατάλληλες σταθερές βάρους και τα  $d$  το στοιχείο μήκους αψίδας. Από άποψης

ανάλυσης εικόνας και όρασης, τα δυο μηχανικά βάρη  $\alpha$  και  $\beta$  θα μπορούσαν να θεωρηθούν σαν η οπτική ισορροπία ανάμεσα στην ευθεία και την καμπύλη. Υπό κανονικές οριακές συνθήκες, οι

Μπίρκοφ και Ντε Μπουρ ονόμασαν τις καμπύλες ισοζυγίου του  $e[\gamma]$  μη-γραμμικές σφήνες στη θεωρία προσέγγισης [27]. Η ενέργεια καμπύλης υδροστατικού σχήματος του Έουλερ αναφέρθηκε για πρώτη φορά στην όραση υπολογιστή από τον Μάμφορτ [222] σαν ένας καλός υποψήφιος για την οπτική παρέμβαση και εφαρμόστηκε πρόσφατα σε έργα επεξεργασίας εικόνας από τους Μασνού και Μορέλ [214], Τσαν, Κανγκ και Σιεν [61], Εσέντογλου και Σιεν [116] και Σια [271].

Επιβάλλοντας την ενέργεια καμπύλης υδροστατικού σχήματος σε κάθε ξεχωριστό επίπεδο καμπύλης

$\gamma_\lambda$  και ζυγίζοντάς τα ομοιόμορφα κάποιος εξασφαλίζει μια μέτρηση κανονικότητας στο διάστημα επιπέδου-συνόλου:

$$E[\Gamma_u] = \int_{-\infty}^{\infty} e[\gamma_\lambda] d\lambda.$$



Για μια λεία εικόνα  $u$ , εφαρμόζοντας την ίδια τεχνική όπως και για την παραγωγή της φόρμουλας συν-περιοχής, κάποιος αντιλαμβάνεται ότι η μέτρηση κανονικότητας θα μπορούσε να εκφραστεί άμεσα από την ίδια την εικόνα:

$$E[\Gamma_u] = E[u] = \int_{\Omega} \left( \alpha + \beta \left( \nabla \cdot \left[ \frac{\nabla u}{|\nabla u|} \right] \right)^2 \right) |\nabla u| dx.$$

### 3.5.6 Στατιστική Επιπέδων Συνόλων Φυσικών Εικόνων

Τώρα παρουσιάζουμε με συντομία την αξιοσημείωτη δουλειά των Γκουσώ και Μορέλ [142] στις στατιστικές των σωρευτικών επιπέδων συνόλων φυσικών εικόνων και τις σημαντικές υπόνοιες στη μοντελοποίηση εικόνων. Θα αλλάξουμε ελαφρώς την παρουσίαση τους (π.χ., ιδέες και εγγραφές) για να ταιριάζει με τη ροή του παρόντος βιβλίου.

Για κάθε επίπεδο έντασης  $\lambda$  και ένα θετικό άλμα  $\Delta\lambda = h \ll 1$ , ορίζουμε

$$L_{\lambda,h}(u) = F_{\lambda+h}(u) \setminus F_{\lambda}(u) = \{x \in \Omega \mid u(x) \in (\lambda, \lambda + h]\}.$$

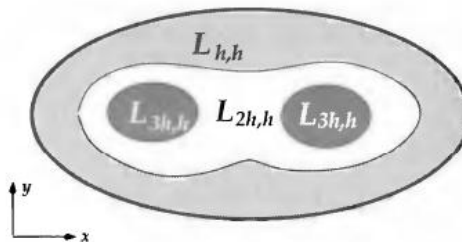
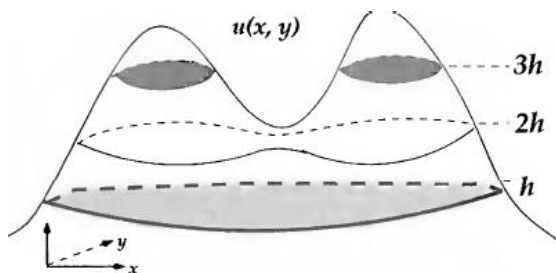
Το  $L_{\lambda,h}$  θα ονομάζεται ένας κλάδος  $h$  του  $u$ . Τότε για οποιοδήποτε  $h$ , ο τομέας εικόνας  $\Omega$  μπορεί να τμηματοποιηθεί σε μη-αλληλοκαλυπτόμενους κλάδους  $h$  για οποιαδήποτε δεδομένη εικόνα  $u$ :

$$\Omega = \bigcup_{n \in \mathbb{Z}} L_{nh,h}(u).$$

Οποιαδήποτε συνδεδεμένη συνιστώσα ενός κλάδου- $h$   $L_{nh,h}(u)$  θα ονομαστεί ένα φυλλάδιο- $h$  του  $u$  (βλέπε Εικόνα 3.5).

Έστω ότι το  $a$  υποδηλώνει την μη-αρνητική πραγματική μεταβλητή για την περιοχή ενός γενικού φυλλαδίου. Ορίζουμε

$$\rho(a \mid h) da = E_u \left[ \frac{\# \text{ of } h\text{-leaflets whose areas are between } (a, a + da)}{\text{total } \# \text{ of } h\text{-leaflets}} \right].$$



Εικόνα

3.5. Ένα παράδειγμα κλάδων  $L_{\lambda, \Delta\lambda}$  με  $\Delta\lambda = h$  και  $\lambda = nh$ . Ο κλάδος  $L_{3h, h}$  περιέχει δύο φυλλαδιά.

Εδώ το  $u$  αντιμετωπίζεται ως ένα τυχαίο πεδίο (και συχνά θεωρείται αναλλοίωτο υπό την αλλαγή γκρίζας κλίμακας:  $u \rightarrow u + \lambda$ ), και το  $E_u$  υποδηλώνει το αντίστοιχο μέσο σύνολο. Για εικόνες αληθινής φύσης, το  $E_u$  συχνά υπολογίζεται από ένα εμπειρικό μέσο όπως στο [142] κάτω από την υπόθεση εργοδικότητα.

Στατιστική έρευνα στις συστοιχίες φυσικών εικόνων των Γκουσώ και Μορέλ αποκαλύπτει τον ακόλουθο νόμο κλίμακας κοντά στο υπεριώδες όριο (δηλαδή,  $a \rightarrow 0^+$ ):

$$\rho(a | h) = \frac{C(h)}{a^\alpha} \text{ με τον εκθέτη } \alpha \text{ κοντά στο } 2.$$

Εδώ το  $C(h)$  υποδηλώνει μια θετική σταθερά η οποία εξαρτάται μόνο από το  $h$ . Είναι αυτός ο σημαντικός στατιστικός νόμος κλιμάκωσης που επέτρεψε στους Γκουσώ και Μορέλ να καταλήξουν ότι οι φυσικές εικόνες δεν έχουν ακριβώς BV. Παραπέμπουμε τον αναγνώστη στο [142] για περαιτέρω λεπτομέρειες.

### 3.6 Το Μοντέλο Ελεύθερου Ορίου Εικόνας Μάμφορτ-Σια

Οι εικόνες είναι οι δυσδιάστατες οπτικές προβολές του τρισδιάστατου κόσμου, ο οποίος είναι γεμάτος από ξεχωριστά αντικείμενα. Κάτω από μια πηγή κοινού φωτός, η εικόνα (ή απεικονισμένη ένταση φωτός) κάθε αντικείμενου χαρακτηρίζεται απόλυτα από το υλικό επιφάνειάς του (δηλαδή αντικατοπτρισμός) καθώς και τοπογραφικές ιδιότητες (δηλαδή γεωμετρία σχήματος). Εφόσον πολύ συχνά κάθε τρισδιάστατο αντικείμενο έχει περίπου ομοιόμορφες ιδιότητες επιφάνειας (όπως τα χέρια ενός ατόμου ή ένα πορτοκάλι), η εικόνα του είναι αντίστοιχα ομοιόμορφη και ξεχωριστή από τις εικόνες άλλων αντικειμένων.

Συνεπώς, με τις λεπτομέρειες υψής επιφάνειας κάθε ξεχωριστού αντικείμενου να παραμελούνται ή να θολώνουν, τέτοιες εικόνες μπορεί να υπολογιστούν σωστά από τμηματικές σταθερές ή λείες λειτουργίες, με κάθε κομμάτι να αντιστοιχεί στην εικόνα ενός ξεχωριστού τρισδιάστατου αντικείμενου. Αυτή είναι η ποιοτική περιγραφή του μοντέλου εικόνας των Μάμφορτ-Σια.

#### 3.6.1 Τμηματικά Σταθερές Μονοδιάστατες Εικόνες: Ανάλυση και Σύνθεση

Αρχίζουμε τη συζήτηση με αυτή την απλούστατη αλλά ενδιαφέρουσα περίπτωση. Ας θεωρήσουμε μια μονοδιάστατη εικόνα  $u(x)$  που καθορίζεται σε ολόκληρο τον πραγματικό άξονα

$$x \in \Omega = \mathbb{R}^1, \text{ με } \lim_{x \rightarrow \pm\infty} u(x) = 0.$$

**Ορισμός 3.19 (Τμηματικά Σταθερή Εικόνα).** Μια μονοδιάστατη εικόνα  $u(x)$  λέγεται ότι είναι τμηματικά σταθερή αν για κάθε μικροστοιχείο  $x_0$ , υπάρχει μια γειτονιά  $|x - x_0| < \varepsilon$  στην οποία

$$u(x) = a + b H(x - x_0)$$

για δυο βαθμωτά  $a$  και  $b$  (εξαρτώμενα από το  $x_0$ ). Εδώ το αντιπροσωπεύει τον κανονικό πίνακα του Χέβισάιτ 0-1:

$$H(s) = \begin{cases} 0, & s < 0, \\ 1, & s \geq 0. \end{cases}$$

Προσέξτε ότι τεχνικά έχει υποτεθεί σιωπηρά ότι η εικόνα  $u$  είναι άμεσα συνεχόμενη, πράγμα το οποίο δεν είναι σημαντικό με την έννοια του Λεπέσκ.

Αν  $b = [u]_{x_0} = u(x_0^+) - u(x_0^-)$  είναι μη-μηδενικό, θα αποκαλείται το άλμα στο  $x_0$ , και  $x_0$  ένα άλμα ή μικροστοιχείου άκρου. Έστω ότι το  $J$  υποδηλώνει στο σύνολο άλματος

$$J = J_u = \{x : [u]_x \neq 0\}.$$

**Θεώρημα 3.20.** Το σύνολο άλματος  $J_u$  οποιασδήποτε τμηματικά σταθερής εικόνας  $u$  είναι κλειστό και απομονωμένο.

Συγκεκριμένα,  $J \cap [A, B]$  είναι ένα πεπερασμένο σύνολο για οποιαδήποτε περιορισμένο διάστημα  $[A, B] \subseteq \mathbb{R}^1$ , και το  $J$  μπορεί να αριθμηθεί από

$$\dots < x_n < x_{n+1} < \dots$$

με  $\lim_{n \rightarrow \pm\infty} x_n = \pm\infty$  αν η ακολουθία είναι άπειρη σε όποια κατεύθυνση.

Με το σύνολο άλματος αριθμημένο όπως πάνω, η εικόνα  $u$  θα μπορούσε να εκφραστεί μοναδικά ως

$$u(x) = \sum_n c_n \chi_{[x_n, x_{n+1})}(x), \quad (3.58)$$

όπου το άθροισμα σε οποιαδήποτε κατεύθυνση θα μπορούσε να είναι πεπερασμένο ή άπειρο σύμφωνα με το  $J_u$ .

Πέρνοντας το γενικοποιημένο ή κατανεμιστικό παράγωγο στο (3.58) κάποιος λαμβάνει

$$Du = \sum_n b_n \delta(x - x_n), \quad (3.59)$$

όπου  $b_n = c_n - c_{n-1} = [u]_{x_n} \neq 0$  είναι τα άλματα, και  $\delta$  είναι το σημείο μέτρησης του Ντιράκ.

Η τελευταία εξίσωση δίνει ένα αποτελεσματικό τρόπο να αναπαραστήσουμε μια μονοδιάστατη τμηματικά σταθερή (και άμεσα συνεχόμενη) εικόνα. Με άλλα λόγια, αρκεί μόνο να κωδικοποιήσουμε τις διευθύνσεις και αξίες όλων των αλμάτων:

$$\{(x_n, b_n) : n\} = \{(z, [u]_z) : z \in J_u\},$$

αντί για ολόκληρη τη μονοδιάστατη λειτουργία.

Αυτή η αναπαράσταση είναι πράγματι χωρίς απώλειες καθώς επιτρέπει ένα τέλειο σχήμα αναδόμησης:

$$\tilde{u}(x) := \tilde{u}(0) + \int_{(0,x]} Du = \tilde{u}(0) + \sum_{x_n \in (0,x]} b_n.$$

Εφόσον έχει υποθεθεί ότι  $u(\pm\infty) = 0$ , η άγνωστη σταθερά  $\tilde{u}(0)$  δίνεται κατηγορηματικά από το

$$\tilde{u}(0) = - \lim_{K \rightarrow +\infty} \sum_{x_n \in (0,K]} b_n.$$

$$\tilde{u} = u.$$

Τότε η σύνθεση πρέπει να είναι ακριβής:

### 3.6.2 Τμηματικά Λείες Μονοδιάστατες Εικόνες: Αναπαράσταση Πρώτης Τάξης

Για τεχνική απλότητα όλες οι εικόνες σε αυτή την υποενότητα θεωρούνται ως συμπαγώς στηριζόμενες.

**Ορισμός 3.21 (Τμηματικά Λείες Εικόνες).** Μια άμεσα συνεχόμενη εικόνα  $u(x)$  στο  $\mathbb{R}^1$  λέγεται ότι είναι τμηματικά λεία αν σε κάθε μικροστοιχείο  $x_0$ , υπάρχει μια γειτονιά  $|x - x_0| < \varepsilon$  στην οποία

$$u(x) = a(x) + b(x)H(x - x_0),$$

όπου  $a(x)$  και  $b(x)$  είναι λειτουργίες  $C^1$  στη γειτονιά.

Σε αντίθεση με την περίπτωση τμηματικής σταθερής, γενικά  $a(x)$  και  $b(x)$  δεν είναι μοναδικά.

Παρόλα αυτά,  $b(x_0)$  και  $b'(x_0)$  είναι όντως μοναδικά εφόσον μπορούν να χαρακτηρίζονται απόλυτα από άποψη της ίδιας της εικόνας:

$$\begin{aligned} b(x_0) &= [u]_{x_0} = u(x_0^+) - u(x_0^-), \\ b'(x_0) &= [u']_{x_0} = u'(x_0^+) - u'(x_0^-). \end{aligned}$$

Ορίζουμε το σύνολο άλματος  $J_u$  ως

$$J_u = \{z \in \mathbb{R} : [u]_z \neq 0\}.$$

Όπως στην περίπτωση της τμηματικής σταθεράς, το  $J_u$  είναι κλειστό και απομονωμένο και μπορεί να είναι πεπερασμένα αριθμημένο ως (λόγω της υπόθεσης συμπαγούς στήριξης)

$$x_1 < x_2 < \dots < x_N.$$

Ορίζουμε  $u_n^\pm = u(x_n^\pm)$  και  $b_n = [u]_n = u_n^+ - u_n^-$  για όλα τα  $1 \leq n \leq N$ . Επιπλέον, ορίζουμε  $x_0 = -\infty$  και  $x_{N+1} = +\infty$ . Τότε σε κάθε διάστημα  $(x_n, x_{n+1})$  με  $0 \leq n \leq N$ , το  $g_n(x) = u'(x)$  ανήκει στο  $L^2(x_n, x_{n+1}) \cap L^1(x_n, x_{n+1})$ .

Μια καλή αναπαράσταση μιας τέτοιας εικόνας δίνεται από το ακόλουθο σύνολο δεδομένων:

$$g_0(x) \text{ and } (x_n, b_n, g_n(x))_{\mu\epsilon} 1 \leq n \leq N. \quad (3.60)$$

Προσέξτε ότι τόσο το  $g_0$  όσο και το  $g_N$  είναι συμπαγώς στηριζόμενα. Αυτή η αναπαράσταση θα μπορούσε να οδηγήσει σε αποτελεσματικά σχήματα κωδικοποίησης όταν η συνδεδεμένη εικόνα  $u$  είναι σχεδόν τμηματικά σταθερή, με άλλα λόγια όταν

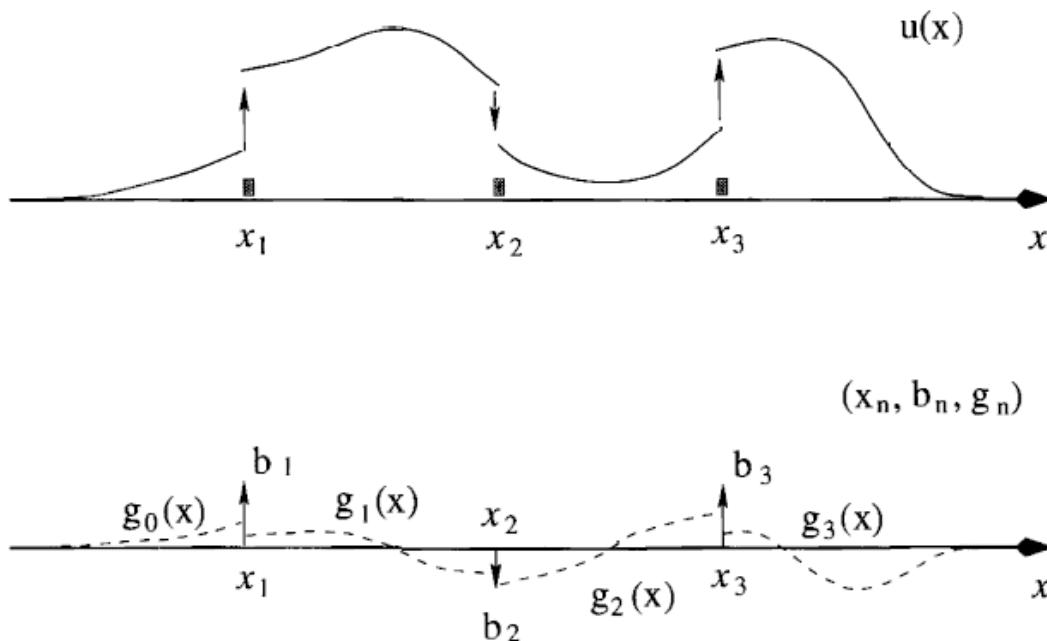
$$\max_{0 \leq n \leq N} \int_{(x_n, x_{n+1})} |g_n(x)| dx \ll 1.$$

Σε αυτή την περίπτωση τα  $g_n$  θα μπορούσαν να κωδικοποιηθούν χρησιμοποιώντας πολύ χαμηλούς ρυθμούς ψηφίων. Στο ακραίο σενάριο όταν τα ψηφία δεν κατανέμονται καθόλου στα  $g_n$ , είναι ισοδύναμο με τον υπολογισμό της αρχικής εικόνας  $u$  από μια τμηματικά σταθερή εικόνα. Γενικά τα  $g_n$  μπορούν να υπολογιστούν σωστά από πολυώνυμα ή σφήνες κατώτερης τάξης. Η Εικόνα 3.6 δείχνει ένα τυπικό παράδειγμα για τονίσει τις υποκείμενες ιδέες.

Η αναπαράσταση (3.60) είναι χωρίς απώλειες καθώς επιτρέπει ένα τέλειο σχήμα αναδόμησης που δίνεται από

$$\tilde{u}(x) = \sum_{x_n \in (-\infty, x]} b_n + \int_{-\infty}^x \sum_0^N g_n(y) \chi_{(x_n, x_{n+1})}(y) dy.$$

Εφόσον τόσο το  $\tilde{u}$  όσο και το  $u$  είναι μηδενικά κοντά στο  $-\infty$  και άμεσα συνεχιζόμενα με πανομοιότυπα κατανεμιστικά παράγωγα  $D\tilde{u} = Du$ , πράγματι πρέπει να ισχύει η τέλεια σύνθεση:  $\tilde{u}(x) \equiv u(x)$ .



**Εικόνα 3.6.** Η αναπαράσταση  $(x_n, b_n, g_n)$  ενός (συμπαγώς στηριζόμενου) τμηματικά λείου σήματος  $u$ . Σε λείες περιοχές όπου το σήμα  $u$  ποικίλει αργά, το  $g_n$  είναι συχνά μικρό και μόνο μερικά ψηφία αρκούν για τους κωδικοποιήσουμε.

### 3.6.3 Τμηματικά Λείες Μονοδιάστατες Εικόνες: Η Αναπαράσταση Poisson

Άλλη μια ίσως πιο ενδιαφέρουσα αναπαράσταση που προκύπτει από την προηγούμενη ενότητα είναι η Αναπαράσταση του Poisson όπως παρακινείται από τη θεωρία των γραμμικών ελλειπτικών ΛΣΕ.

Ας υποθέσουμε ότι το  $u$  είναι Σοπόλεβ  $H^2$  σε κάθε διάστημα  $(x_n, x_{n+1})$  ανάμεσα σε δυο γειτονικά σημεία άλματος. Με άλλα λόγια,  $f_n(x) = u''(x) \in L^2(x_n, x_{n+1})$ . Τότε η αναπαράσταση Poisson του  $u$  αναφέρεται στο ακόλουθο σύνολο δεδομένων:

$$f_0(x) \text{ and } (x_n, u_n^+, u_n^-, f_n(x)) \text{ with } 1 \leq n \leq N. \quad (3.61)$$

Όπως και στην προηγούμενη ενότητα, προσέξτε ότι τόσο το  $f_0$  όσο και το  $f_N$  είναι συμπαγώς στηριζόμενα.

Αυτή η αναπαράσταση θα μπορούσε να είναι πολύ αποτελεσματική όταν η εικόνα στόχος  $u$  είναι σχεδόν γραμμική σε κάθε διάστημα  $(x_n, x_{n+1})$  στην οποία περίπτωση κάποιος έχει

$$\max_{0 \leq n \leq N} \int_{(x_n, x_{n+1})} |f_n(x)|^2 dx \ll 1.$$

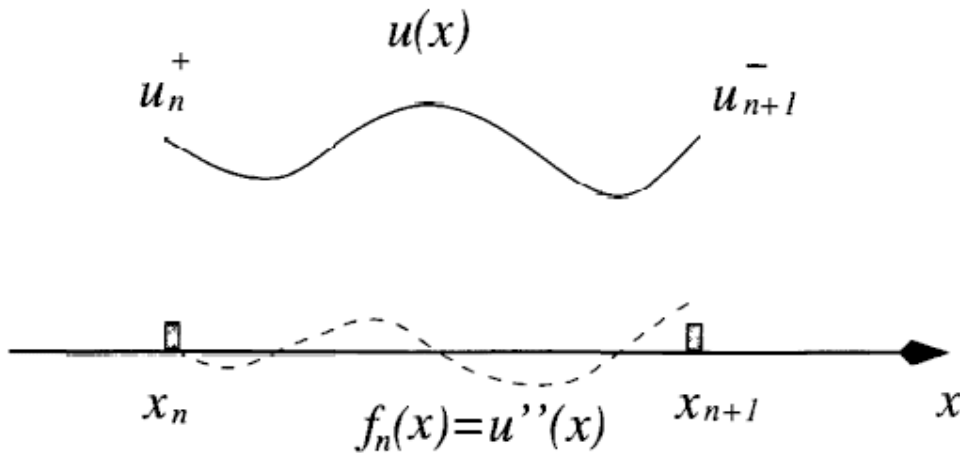
Συνεπώς, όλα τα  $f_n$  μπορούν να κωδικοποιηθούν χρησιμοποιώντας πολύ χαμηλά ψηφία. Όταν τα  $f_n$  έχουν αξαλυφθεί όλα στα μηδεν, είναι ισάξια στον υπολογισμό της αρχικής εικόνας  $u$  από μια τμηματική γραμμική λειτουργία.

Η αναπαράσταση Poisson (3.61) είναι επίσης χωρίς απώλειες εφόσον μια τέλεια αναδόμηση μπορεί να επιτευχθεί λύνοντας τις εξισώσεις του Poisson με τα οριακά δεδομένα Ντίριχλετ: σε κάθε διάστημα  $x \in (x_n, x_{n+1})$ ,

$$\tilde{u}''(x) = f_n(x) \text{ with } \tilde{u}(x_n^+) = u_n^+ \text{ and } \tilde{u}(x_{n+1}^-) = u_{n+1}^-. \quad (3.62)$$

Όλα τα απαραίτητα δεδομένα για να θέσουμε αυτές τις εξισώσεις είναι διαθέσιμα από την αναπαράσταση του Poisson (3.61). Από τη θεωρία γραμμικών ελλειπτικών εξισώσεων, οι αδύναμες λύσεις  $H^1$  υπάρχουν και είναι μοναδικές για κάθε σύστημα. Έτσι αφού κολλήσουμε όλα τα κομμάτια

$\tilde{u}$  σύμφωνα με την άμεσα συνεχιζόμενη υπόθεση, κάποιος πρέπει να έχει την τέλεια αναδόμηση:  
 $\tilde{u}(x) \equiv u(x)$  (βλέπε Εικόνα 3.7).



**Εικόνα 3.7.** Η αναπαράσταση Poisson ενός τμηματικά λείου σήματος  $u$ . Εδώ φαίνεται μόνο ένα μοναδικό αναπαραστατικό διάστημα. Το σήμα  $u$  αναπαριστάται από τις δυο οριακές αξίες  $u_n^+$  και  $u_{n+1}^-$  και το παράγωγό του δευτερεύουσας τάξης  $f = u''$  (αντιστοιχεί στην πηγή κατανομής στον ηλεκτρομαγνητισμό). Η αναδόμηση του  $u$  στο διάστημα τότε αντιστοιχεί στην επίλυση της εξίσωσης του Poisson (3.62). Το πλεονέκτημα μιας τέτοιας αναπαράστασης είναι ότι το  $f$  είναι συχνά μικρό (ακριβώς όπως οι συντελεστές κυματιδίων) για λεία σήματα και απαιτεί λιγότερα ψηφία σε σύγκριση με το  $u$ .

### 3.6.4 Τμηματικά Λείες Δυσδιάστατες Εικόνες

Στις δυο διαστάσεις, η κύρια πρόκληση προκύπτει από τη γεωμετρική πολυπλοκότητα του συνόλου άλματος  $J_u$ . Ακολουθώντας την αρχική συζήτηση της παρορσασ ενότητας, το σύνολο άλματος της εικόνας θα έπρεπε να παγιδεύει περίπου τα προβλημένα όρια των τρισδιάστατων αντικειμένων. Λογω της αδιαφάνειας και έγκλεισης, η συλλεκτική συμπεριφορά του συνόλου άλματος θα μπορούσε να είναι πού αναμεμυγμένο. Γωνίες, συνδέσεις  $T$ , ατελή άκρα, και ταλαντωτικά όρια είναι μόνο μερικλα από τα τυπικά παραδείγματα.

Τεράστιες προκλήσεις υπάρχουν ακόμη, τόσο από άποψης θεωρητικής ανάλυσης, όσο και από άποψης υπολογιστικών προσεγγίσεων, από τις οποίες κάποιος θα μπορούσε να πάρει μια ιδέα τόσο από την επόμενη υποενότητα όσο και από το τελευταίο κεφάλαιο. Στη συνέχεια, έχει γίνει μια απαραίτητη απλοποίηση για να διευκολύνει τη συζήτηση.

**Ορισμός 3.22 (Τμηματοποίηση Λιπσχιτζ ενός Τομέα).** Έστω ότι  $\Omega \subseteq \mathbb{R}^2$  είναι ένας τομέας εικόνας στόχου, ο οποίος είναι περιορισμένος, ανοιχτός. Και Λιπσχιτζ (συχνά τετράγωνο). Μια τμηματοποίηση Λιπσχιτζ του  $\Omega$  αναφέρεται σε ένα πεπερασμένο σύνολο τμηματοποίησης



$$\Omega = \bigcup_{n=1}^N \Omega_n \cup \Gamma \quad (3.63)$$

το οποίο ικανοποιεί τα ακόλουθα:

1. Κάθε  $\Omega_n$  είναι ένας συνδεδεμένος ανοιχτός τομέας.
2. Το  $\Gamma$  είναι σχετικά κλειστό στο  $\Omega$  και έχει ένα πεπερασμένο μέτρο μονοδιάστατου Χάουσοντορφ  $\mathcal{H}^1(\Gamma) < \infty$ .

**Θεώρημα 3.23.** Ας υποθέσουμε ότι μια τμηματοποίηση Λιπσχιτζ του  $\Omega$  δίνεται από το (3.63). Για κάθε υποσύνολο  $X$  του  $\Omega$ , έστω ότι το  $\partial_\Omega X$  υποδηλώνει το σχετικό τοπολογικό όριο στο  $\Omega$ . Τότε

$$(1) \partial_\Omega \Gamma = \Gamma;$$

$$(2) \partial_\Omega \Omega_n \subseteq \Gamma \text{ for } n = 1 : N.$$

$$(3) \Gamma = \bigcup_{n=1}^N \partial_\Omega \Omega_n.$$

**Απόδειξη.** (1) είναι αλήθεια εφόσον το εσωτερικό του  $\Gamma$  πρέπει να είναι κενό, αλλιώς η μέτρηση μονοδιάστατου Χάουσοντορφ θα ήταν άπειρη. (2) πρέπει να ισχύει εφόσον το

$$\partial_\Omega \Omega_n \cap \partial_\Omega \Omega_m \text{ πρέπει να είναι κενό για κάθε ζευγάρι } (n, m).$$

(2) δηλώνει ότι

$$\bigcup_{n=1}^N \partial_\Omega \Omega_n \subseteq \Gamma.$$

Από την άλλη, για κάθε μικροστοιχείο  $z \in \Gamma = \partial_\Omega \Gamma$ , υπάρχει τουλάχιστο ,μια ακολουθία μικροστοιχείων  $(x_k)_{k=1}^\infty \subseteq \Omega \setminus \Gamma$ , έτσι ώστε  $x_k \rightarrow z$  ως  $k \rightarrow \infty$ . Εφόσον το  $N$  είναι πεπερασμένο, πρέπει να υπάρχει συγκεκριμένο  $m$ , έτσι ώστε η ακολουθία να επισκέπτεται το  $\Omega_m$  άπειρα συχνά. Αυτό δηλώνει ότι  $z \in \partial_\Omega \Omega_m$  και συνεπώς,

$$\Gamma \subseteq \bigcup_{n=1}^N \partial_\Omega \Omega_n.$$

Έτσι το (3) πρέπει να είναι επίσης αλήθεια.

Αυτό το θεώρημα επιβεβαιώνει ότι για μια τμηματοποίηση Λίπσχιτζ, το σύνολο  $\Gamma$  μοντελοποιεί όντως τα όρια διαφορετικών «αντικειμένων».

**Ορισμός 3.24.** Μια εικόνα  $u$  στο  $\Omega$  λέγεται ότι είναι τμηματικά λεία αν υπάρχει μια τμηματοποίηση Λίπσχιτζ του  $\Omega$  όπως δίνεται στο (3.63) και  $N$  λείες λειτουργίες στο  $\Omega$ :  $U_1(x), U_2(x), \dots, U_N(x)$ , τέτοιες που

$$u(x) = \sum_{n=1}^N U_n(x) \chi_n(x), \quad x \in \Omega \setminus \Gamma,$$

όπου  $\chi_n(x) = \chi_{\Omega_n}(x)$  είναι οι δείκτες.

Προσέξτε ότι ο ορισμός δεν διευκρινίζει τις αξίες εικόνας κατά μήκος του οριακού συνόλου  $\Gamma$ . Αυτό δουλεύει σωστά από μετρικής – θεωρητικής άποψης εφόσον  $\Gamma$  είναι ένα δυσδιάστατο μηδενικό σύνολο Λεπέσκ.

Όπως και στη μονοδιάστατη περίπτωση, εδώ ο όρος «λείο» θα μπορούσε να αναφέρεται τουλάχιστο στην κλασσική κανονικότητα  $C^1$  ή πιο γενικά την κανονικότητα του Σοπόλεβ  $H^1$ . Κάθε  $u_n(x) = U_n(x) \chi_n(x)$  μοντελοποιεί στο σχετικό ομοιογενές εσωτερικό ενός μοναδικού «αντικειμένου».

Παρόμοια με τη μονοδιάστατη περίπτωση, μια αναπαράσταση Poisson ή κωδικοποιητής θα μπορούσαν να δομηθούν ως ακολούθως. Ας υποθέσουμε ότι κάθε «αντικείμενο»  $u_n$  ανήκει στο  $H^2(\Omega_n)$ . Τότε η Λαπλασιανή  $f_n(x) = \Delta u_n$  ανήκει στο  $L^2(\Omega_n)$ . Στην περίπτωση που κάθε  $u_n$  είναι σχεδόν αρμονικό στο  $\Omega_n$ , κάποιος έχει

$$\sum_{n=1}^N \int_{\Omega_n} |f_n(x)|^2 dx \ll 1,$$

και συνεπώς, μόνο ρυθμοί χαμηλων ψηφίων χρειάζονται για την κωδικοποίηση των  $f_n$ .

Από την άλλη, σύμφωνα με τη θεωρία των διαστημάτων του Σοπόλεβ, εφόσον κάθε  $\Omega_n$  υποτίθεται ότι είναι ένας τομέας Λίπσχιτζ, το ίχνος του  $u_n(x)$  κόντρα στο  $\gamma_n = \partial\Omega_n$  προσδιορίζεται σωστά. Ας υποδηλωθεί από το  $\phi_n(x)$ , προσδιορισμένο κατά μήκος του  $\gamma_n$ . Προσέξτε ότι στο  $R^2$ ,

$$\bigcup_{n=1}^N \gamma_n = \partial\Omega \cup \Gamma.$$

Τότε μια αναπαράσταση Poisson για μια τέτοια τμηματική λεία εικόνα  $u$  αποτελείται από το ακόλουθο σύνολο δεδομένων:

$$\{(\gamma_n, \phi_n(x), f_n(x)) : 1 \leq n \leq N\}. \quad (3.64)$$

Προσέξτε ότι τα  $\phi_n(x)$  είναι στην ουσία μονοδιάστατες λειτουργίες (προσδιορισμένα κατά μήκος των  $\gamma_n$ ). Σε σύγκριση με την άμεση κωδικοποίηση των  $u_n$ , τα  $f_n$  είναι συχνά μικρά και απαιτούν μόνο χαμηλούς ρυθμούς ψηφίων κωδικοποίησης για σχεδόν αρμονικά αντικείμενα.

Η αναπαράσταση του Poisson (3.64) είναι χωρίς απώλειες αφού επιτρέπει ένα τέλει σχήμα

αναδόμησης με την μετρική-θεωρητική έννοια. Με άλλα λόγια στο  $\Omega \setminus \Gamma$ , η αρχική εικόνα  $u$  θα μπορούσε να είναι τέλεια αναδομημένη. Αυτό γίνεται αντιληπτό από τον επιλυτή του Poisson σε κάθε

συνιστώσα Λιπσχιτζ  $\Omega_n$ :

$$\Delta \tilde{u}(x) = f_n(x) \text{ for } x \in \Omega_n \text{ and } \tilde{u}(y) = \phi_n(y) \text{ for } y \in \gamma_n.$$

Σύμφωνα με το θεώρημα ύπαρξης και μοναδικότητας για γραμμικές ελλειπτικές εξισώσεις με δεδομένα Νίριχλετ στους τομείς Λιπσχιτζ, η τέλεια αναδόμηση είναι πράγματι αλήθεια:  $\tilde{u}(x) \equiv u(x)$  για κάθε  $x \in \Omega \setminus \Gamma$ .

Μια αρμονική αναπαράσταση θα αποτελείτο μόνο από τα ακόλουθα μονοδιάστατα δεδομένα:

$$\{(\gamma_n, \phi_n(x)) : 1 \leq n \leq N\} \quad (3.65)$$

με όλα τα  $f_n$  ρυθμισμένα στο μηδέν. Συνεπώς, η αρμονική αναπαράσταση είναι με απώλειες και υπολογισμό, το απαραίτητο τίμημα που πρέπει να πληρωθεί που λόγω της υψηλής συμπίεσης δεδομένων.

### 3.6.5 Το Μοντέλο Μάμφορτ –Σια

Ένα κρίσιμο θέμα παραμένει ακόμη για τις πραγματικές εφαρμογές και υπολογισμούς: δεδομένης μιας τμηματικής λείας εικόνας  $u$ , πώς θα μπορούσε κάποιος να εξάγει τη συνδεδεμένη της αποσύνθεση αντικειμένου – οριού;

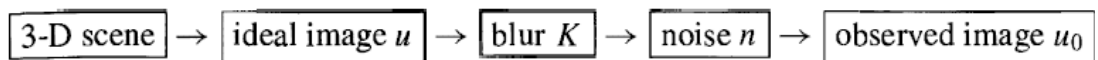
Όπως δηλώθηκε πιο πάνω, οι εικόνες είναι οι οπτικές προβολές τρισδιάστατων σκηνών γεμάτων ξεχωριστών αντικειμένων. Έτσι η διαδικασία εξαγωγής είναι ένα αντίστροφο πρόβλημα που στοχεύει στην επανάκτηση αντικειμένων από εικόνες, ή πιο κατάλληλα, δυσδιάστατη απασχόληση περιοχών ανεξάρτητων αντικειμένων από εικόνες. Μια τέτοια διαδικασία αποκαλείται συχνά το πρόβλημα τμηματοποίησης.

Ως ένα αντίστροφο πρόβλημα, η τμηματοποίηση επιπυγχάνεται αδιάλλακτα από δυο ουσιαστικά ισοδύναμες μεθόδους: στατιστικός υπολογισμός μέσω μέγιστης πιθανότητας (ΜΠ) ή μέγιστες εκ των υστέρων πιθανότητες (ΜΕΠ) και ο προσδιοριστικός υπολογισμός μέσω μεταβλητής βελτιστοποίησης βασισμένης στην ενέργεια. Το πασίγνωστο μοντέλο τμηματοποίησης [226] των Μάμφορτ και Σια ανήκει στη δεύτερη κατηγορία και είναι στενά εμπνευσμένο από και συνδεδεμένο με νωρίτερα στατιστικά μοντέλα, για παράδειγμα, από τους Γκέμαν και Γκέμαν [130] και τους Μπλέικ και Ζίσσερμαν [29].

Ακολούθως, κάνουμε μια σύντομη εισαγωγή στο μοντέλο εικόνας Μάμφορτ-Σια και περισσότερες πληροφορίες σχετικά με της εφαρμογές και τον υπολογισμό του θα εμφανιστούν στο τελευταίο κεφάλαιο.

Έστω ότι  $u_0$  υποδηλώνει μια αποκτημένη εικόνα στον τομέα  $\Omega$ . Στις πλείστες εφαρμογές, είναι ακατάλληλο να υποθέσουμε ευθέως ότι το  $u_0$  είναι τμηματικά λείο, ακόμη και αν η τρισδιάστατη σκηνή από την οποία εξασφαλίστηκε το  $u_0$  αποτελείται από ανεξάρτητα αντικείμενα με λείες διαφοροποιημένες επιφάνειες και ιδιότητες συντελεστή ανάκλασης. Πρώτα, κατά τη διάρκεια της επεξεργασίας εικόνων, ο θόρυβος είναι συχνά αναπόφευκτος καθώς προκαλείται από διάφορους τυχαίους παράγοντες, όπως τα κινούμενα ανθρώπινα σώματα ή όργανα (σαν την καρδιά), στην ιατρική εικονοποίηση, ανομοιογένεια χωρικών μέσων όπως νερό για εικόνες υποθαλάσσιες και φυσικές βιοχημικές αντιδράσεις στα κυτταρικά, νευρικά δίκτυα του συστήματος ανθρώπινης όρασης. Δεύτερον, όπως στην αστρονομική εικονοποίηση ή φωτογραφίες από σπορ, το θολωμένο εφέ που προκαλείται από τοπικά μέσα παρέμβασης ή η σχέση κίνησης είναι επίσης συχνά παρόν.

Τέτοια πολυπλοκότητα πραγματικών περιβάλλοντων έχουν εμπνεύσει πολλά πετυχημένα μοντέλα. Μια από τις πιο πασίγνωστες προσεγγίσεις είναι να διαχωρίσουμε τη διαδικασία απεικόνισης σε διαφορετικά στρώματα, για παράδειγμα,



Έτσι για παράδειγμα, υποθέτοντας το θόρυβο ως προσθετικό, κάποιος έχει

$$u_0 = n + K[u], \quad (3.66),$$

όπου ο θολωμένος τελεστής μπορεί να είναι είτε γραμμικός είτε μη γραμμικός.

Είναι η εικόνα  $u$  που μπορεί να υποτεθεί τμηματικά λεία, τουλάχιστον ως ένας καλός υπολογισμός όταν η τρισδιάστατη σκηνή αποτελείται κυρίως από αντικείμενα χωρίς υφή. Η αποστολή της τμηματοποίησης εικόνων είναι να εξάγει ή να υπολογίσει τις πληροφορίες αντικειμένου-ορίου από ένα μοναδική υποβιβασμένη παρατήρηση  $u_0$ .

Τώρα επεξηγούμε τα υλικά κλειδί τις αντικειμενικής λειτουργίας  $E[u, \Gamma|u_0]$  στην αποφαση τμηματοποίησης των Μάμορτ και Σια. Ας υποθέσουμε ότι η ιδανική εικόνα  $u$  είναι δομημένη πάνω στην τμηματοποίηση του Λίπσχιτς

$$\Omega = \bigcup_{m=1}^M \Omega_m \cup \Gamma,$$

όπως προσδιορίζεται στην προηγούμενη ενότητα, με τις λείες συνιστώσες

$$u_m = u|_{\Omega_m} \in H^1(\Omega_m), \quad m = 1 : M.$$

Σε μια πρόσθετη τάση (η οποία είναι επαρκώς υποστηριγμένη από την άποψη της πιθανότητας), η κανονικότητα ενέργειας του  $\mathbf{u}$  καθορίζεται ως

$$E[u, \Gamma] = E[\Gamma] + E[u|\Gamma].$$

Το σύνολο άκρου  $\Gamma$  μπορεί να αξιολογηθεί φυσικά από κατά μήκος, ή πιο γενικά, η μέτρηση μονοδιάστατου Χάουσοντορφ  $E[\Gamma] = \mathcal{H}^1(\Gamma)$ . Από την άλλη, όλα τα βύσματα  $(u_m)_{m=1}^M$  μπορούν να αξιολογηθούν από τις μονάδες Σοπόμεβ:

$$E[u|\Gamma] \propto \sum_{m=1}^M E[u_m|\Omega_m] = \sum_{m=1}^M \int_{\Omega_m} |\nabla u_m|^2 dx = \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx.$$

Σε συνδυασμό, η ενέργεια κανονικότητας μιας ιδανικής τμηματικής εικόνας  $\mathbf{u}$  με ένα σύνολο άκρων  $\Gamma$  δίνεται από

$$E[u, \Gamma] = \alpha \mathcal{H}^1(\Gamma) + \beta \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx, \quad (3.67)$$

με κατάλληλα σταθερά βάρη  $\alpha$  και  $\beta$  για να μοντελοποιήσουν τις αντίστοιχες ευαισθησίες.

Τέλος, ας υποθέσουμε ότι ο πρόσθετος θόρυβος  $\mathbf{n}$  στο μοντέλο απεικόνισης (3.66) είναι ομογενές Γκαουσιανός λευκός θόρυβος με μεταβλητή  $\sigma^2$ . Τότε η μεταβλητή θα μπορούσε να υπολογιστεί σωστά από τον εμπειρικό εκτιμητή

$$\sigma^2 \simeq \frac{1}{|\Omega|} \int_{\Omega} (u_0 - K[u])^2 dx, \quad (3.68)$$

όπου το  $|\Omega|$  υποδηλώνει την περιοχή Λεπέσκ, δεδομένου ότι το  $\mathbf{u}$  είναι γνωστό.

Τότε ο εκτιμητής τμηματοποίησης Μάμφορτ-Σια είναι να

μειώσει  $E[u, \Gamma]$  υποκείμενο στον περιορισμό θυρύβου (3.68).

Εισάγοντας ένα πολλαπλασιαστή Λανγκραντζ  $\lambda$  για τον περιορισμό, το μοντέλο Μάμφορτ-Σια είναι να μειώσει την μοναδική λειτουργική αμεροληψία έχοντας μια παρατήρηση  $u_0$ :

$$E[u, \Gamma | u_0] = \alpha \mathcal{H}^1(\Gamma) + \beta \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx + \lambda \int_{\Omega} (u_0 - K[u])^2 dx. \quad (3.69)$$

Με τον τελευταίο τρίτο όρο να συμβολίζεται από  $E[u_0 | u, \Gamma]$ , γίνεται απλά

$$E[u, \Gamma | u_0] = E[u, \Gamma] + E[u_0 | u, \Gamma],$$

που δείχνει ξεκάθαρα την Μπεϋζιανή λογική του μοντέλου Μάμφορτ-Σια – την τμηματική λεία εικόνα πριν από το μοντέλο  $E[u, \Gamma]$  και μοντέλο απόκτησης εικόνας  $E[u_0 | u, \Gamma]$  είναι επίσημα συνδυασμένα.

### 3.6.6 Ο Ρόλος των Ειδικών Εικόνων BV

Η πρόκληση για την τμηματικά λεία αναπαράσταση και το μοντέλο τμηματοποίησης Μάμφορτ-Σια είναι το σύνολο άκρου  $\Gamma$ , ή το γεωμετρικό υλικό του μοντέλου.

Ας θεωρήσουμε τη συλλογή όλων αυτών των συνόλων άκρου. Καμία φυσική γραμμική δομή δεν φαίνεται εμφανής για να την κάνει ένα διάστημα Μπάναχ ή Χίλμπερτ, προκαλώντας τη δυσκολία στην αναγνώριση ενός βολικού, επιτρεπτού διαστήματος για τη λειτουργική αντικειμενικότητα του Μάμφορτ-Σια. Ως αποτέλεσμα, τα περισσότερα εργαλεία για τη συμβατική βελτιστοποίηση δεν μπορούν να εφαρμοστούν άμεσα.

Παρόλα αυτά, υπάρχει ένα διάστημα γραμμικής λειτουργίας που είναι πραγματικά αρμόδιο για το μοντέλο Μάμφορτ-Σια. Είναι ένα υποδιάστημα του  $BV(\Omega)$  που αποτελείται από την αποκαλούμενες ξεχωριστές λειτουργίες  $BV(SBV)$  και υποδηλώνεται από το  $SBV(\Omega)$  [7].

Έστω ότι το  $u$  είναι μια επιτρεπτή εικόνα Μάμφορτ-Σια με, για απλότητα, ένα τμηματικό  $C^1$  σύνολο  $\Gamma$  άκρης. Έστω ότι  $\mathcal{H}^1$  υποδηλώνει μια μονοδιάστατη μέτρηση Χάουσαντορφ κατά μήκος του  $\Gamma$ . Τότε, το  $\mathcal{H}^1$  σχεδόν σίγουρα, κάποιος θα μπορούσε να ορίσει μια τμηματική συνεχόμενη κανονική μονάδα φορέα πεδίου  $\nu$  κατά μήκος του  $\Gamma$ . Για κάθε μικροστοιχείο άκρου  $x$  σε μια συνιστώσα του  $\Gamma$ , όπως και στη μονοδιάστατη περίπτωση, ορίζουμε το άλμα  $[u]_x$  ως

$$[u]_x = \lim_{\varepsilon \rightarrow 0^+} u(x + \varepsilon \nu_x) - u(x - \varepsilon \nu_x).$$

Εδώ τα όρια γίνονται κατανοητά με την έννοια της σάρωσης εφόσον κάθε

$u_m = u|_{\Omega_m}$  υποτίθεται στο  $H^1(\Omega_m)$ . Τότε περιορισμένο στο άκρο συνόλου  $\Gamma$  μόνο το,

$$J_u = [u] \nu d\mathcal{H}^1$$

προσδιορίζει μια ανυσματική μέτρηση από, για οποιοδήποτε υποσύνολο Μπόμελ  $\gamma \subseteq \Gamma$ ,

$$J_u(\gamma) := \int_{\gamma} [u]_x \nu_x d\mathcal{H}_x^1$$

Θα μπορούσε να αποδειχτεί εύκολα υπό την υπόθεση της κανονικότητας  $\mathcal{C}^1$  ότι η ανυσματική μέτρηση Ρααδονίου  $Du$  όταν περιορισμένο στο κλειστό σύνολο άκρου  $\Gamma$  είναι ακριβώς το  $J_u$ . Από την άλλη, σε κάθε ανοιχτή συνιστώσα  $\Omega_m$ , με βάση την συνθήκη παραδοχής του μοντέλου Μάμφορτ-Σια,  $u_m \in H^1(\Omega_m)$  κάποιος έχει  $Du = \nabla u \in L^2(\Omega_m)$ . Σε συνδυασμό, έχουμε

$$Du = \nabla u|_{\Omega \setminus \Gamma} + J_u|_{\Gamma}. \quad (3.70)$$

Αυτό το επιχείρημα είναι σχεδόν μη-εριστικό, με την προϋπόθεση ότι η συνθήκη  $\mathcal{C}^1$  είναι αφημένη στην πιο γενική συνθήκη της πεπερασμένης μέτρησης  $\mathcal{H}^1$  όπως στο μοντέλο Μάμφορτ-Σια και ότι τόσο το  $[u]$  όσο και το  $\nu$  είναι κατάλληλα προσδιορισμένα.

Τώρα ρωτάμε πόσο γενική είναι η δομή της αποσύνθεσης των δυο συνιστώσων (3.70) στο  $BV(\Omega)$ . Στην ευλογημένη περίπτωση όταν οποιαδήποτε λειτουργία  $BV$  επέτρεπε τέτοια αναπαράσταση, το διάστημα

$BV(\Omega)$  θα γινόταν τέλει για τη μελέτη του μοντέλου των Μάμφορτ-Σια. Αλλά δεν είναι αυτή η περίπτωση στην πραγματικότητα.

Έστω ότι το  $u \in BV(\Omega)$  είναι μια γενική εικόνα με  $BV$ . Σύμφωνα με την αποσύνθεση των Ρέιτον, Νικόδημο και Λεπέσκ, κάποιος θα μπορούσε να γράψει

$$Du = \nabla u + D_s u,$$

όπου  $\nabla u = Du/dx$  είναι το παράγωγο των Ρέιτον-Νικόδημου και κωδικοποιεί πληροφορίες εικόνας που μπορούν να απαθανατιστούν από τη μέτρηση Λεπέσκ

$dx = dx_1 dx_2$ , εφόσον το  $D_s u$  είναι μοναδικό στη μέτρηση Λεπέσκ και υποστηρίζεται μόνο από ένα μηδενικό σύνολο Λεπέσκ. Γενικά μιλώντας, η διάσταση Χάουσοντορφ αυτού του μηδενικού συνόλου πρέπει να είναι λιγότερη από 2 επίπεδες εικόνες. Είναι αυτή η μοναδική συνιστώσα που περιπλέκει μια γενική εικόνα  $BV$ .

Το μοναδικό κομμάτι του Λεπέσκ θα μπορούσε να αποσυντεθεί περαιτέρω σε

$$D_s u = D_1 u + D_c u, \quad (3.71)$$

όπου, εριστικά μιλώντας, το  $D_1 u$  κωδικοποιεί μονοδιάστατες πληροφορίες που μπορούν να παγιδευτούν από σκορπισμένες μονοδιάστατες μετρήσεις Χάουσοντορφ  $\mathcal{H}^1$  στον τομέα εικόνας, ενώ το

$D_c u$  αποθηκεύει πληροφορίες που υποστηρίζονται από ένα σύνολο των οποίων η διάσταση Χάουσοντορφ είναι αυστηρώς ανάμεσα στο 1 και 2. Ο δείκτης «c» αντιπροσωπεύει τον «Κατανομέα»

(Cantor) εφόσον η λειτουργία Κατανομέα [126]  $f(x)$  είναι ιστορικά το πιο πασίγνωστο παράδειγμα σε μια διάσταση που είναι σχεδόν (με την έννοια της μιας διάστασης του Λεπέσκ) πάντα επίπεδη στο  $[0, 1]$  αλλά μαγικά αυξάνεται με μια συνεχόμενη τάση από  $f(0) = 0$  σε  $f(1) = 1$ . Η αύξηση επιτυγχάνεται στο σύνολο του Κατανομέα, το οποίο έχει μια κλασματική διάσταση. Έτσι το

$D_c u$  αποκαλείται συχνά συνιστώσα μοναδικού Κατανομέα για μια γενική εικόνα  $BV u$ . Χωρίς ένα σαφή προσδιορισμό του συνόλου άκρου  $\Gamma$  όπως στο μοντέλο Μάμφορτ-Σια, πώς θα μπορούσε να ανιχνευθεί ή να προσδιοριστεί κατάλληλα το μονοδιάστατο χαρακτηριστικό  $D_1 u$  για μια γενική εικόνα  $u$ ;

Δεδομένης μιας γκριζας κλίμακας επιπέδου  $\lambda$ , ενός μικροστοιχείου  $z \in \Omega$ , και μιας μικρής ακτίνας  $\rho > 0$ , ορίζουμε ένα υποσύνολο του δίσκου  $B_{z,\rho} = \{x : |x - z| < \rho\}$  από

$$\{u > \lambda\}_{z,\rho} = \{x \in \Omega \cap B_{z,\rho} : u(x) > \lambda\}$$

Ο ίδιος ορισμός επεκτείνεται επίσης στο σήμα  $\cdot$ .

**Ορισμός 3.25 (Ουσιαστικά Όχι Μεγαλύτερο Από ή Όχι Λιγότερο Από).** Το  $u$  λέγεται ότι είναι (Λεπέσκ) ουσιαστικά όχι μεγαλύτερο από το  $\lambda$  στο  $x$  αν

$$\lim_{\rho \rightarrow 0^+} \frac{|\{u > \lambda\}_{x,\rho}|}{|B_{x,\rho}|} = 0$$

όπου το  $|\cdot|$  αντιπροσωπεύει τη δυσδιάστατη μέτρηση στο  $\Omega$ . Αν ισχύει αυτό, γράφουμε

$u \leq_x \lambda$  Ο ίδιος ορισμός και εγγραφή επεκτείνονται φυσικά στο να «είναι όχι μεγαλύτερα από.»

Η υπόνοια είναι ξεκάθαρη: αν το  $u$  είναι ουσιαστικά όχι μεγαλύτερο από το  $\lambda$  στο  $x$  καθώς κάποιος εστιάζει στη γειτονιά, το αντίστροφο γεγονός  $u > \lambda$  θα έπρεπε να γίνει σπανιότερο και σπανιότερο, μέχρι που στο τέλος εξαφανίζεται εντελώς με τη μετρική-θεωρητική έννοια.

**Ορισμός 3.26 (Υψηλότερες και Χαμηλότερες Λειτουργίες).** Δεδομένης μιας μετρήσιμης λειτουργίας  $u$  στο  $\Omega$ , ορίζουμε την υψηλότερη λειτουργία της από το

$$u_+(x) = \inf\{\lambda \in \mathbb{R} : u \leq_x \lambda\}$$



ο υψηλότερος περιορισμός όλων των αξιών ουσιαστικά όχι μεγαλύτερος από το  $u$  στο  $x$ .

Είναι εύκολο να δείξουμε ότι αν  $u \leq_x \lambda$  και  $u \geq_x \mu$ , τότε  $\mu \leq \lambda$ . Συνεπώς, πρέπει να έχουμε  $u_-(x) \leq u_+(x)$  σε οποιοδήποτε μικροστοιχείο  $x$ . Με μια φυσική έννοια, θα μπορούσαμε να σκεφτούμε ότι το  $u$  είναι ουσιαστικά περιορισμένο ανάμεσα στα  $u_-(x)$  και  $u_+(x)$  στο  $x$ . Από την άλλη, τόσο το  $u_+(x)$  όσο και το  $u_-(x)$  μπορούν να επιτευχθούν με την ακόλουθη έννοια. Για κάθε  $\lambda < u_+(x)$  υπάρχει μια ακολουθία  $\rho_k \rightarrow 0$  ( $k \rightarrow \infty$ ) και κάποια σταθερά  $r > 0$  έτσι ώστε

$$\lim_{k \rightarrow \infty} \frac{|\{u > \lambda\}_{x, \rho_k}|}{|B_{x, \rho_k}|} = r > 0$$

Το ίδιο μπορεί να ειπωθεί για το  $u_-(x)$ .

Το σύνολο άλματος  $S_u$  μιας εικόνας  $BV$   $u \in BV(\Omega)$  προσδιορίζεται τότε ως

$$S_u = \{x \in \Omega : u_+(x) > u_-(x)\}. \quad (3.72)$$

Η καθοδήγηση αυτού του ορισμού είναι η απλή υπόνοια ότι ένα σημείο άλματος είναι ένα μικροστοιχείο όπου συμβαίνει ένα ουσιαστικό άλμα. Έτσι εισάγεται το  $S_u$  σαν υποψήφιο για το σύνολο άκρου

$\Gamma$  στο μοντέλο Μάμφορτ-Σια. Προσέξτε, παρόλα αυτά, ότι ο προσδιορισμός του  $S_u$  είναι απόλυτα ενδογενείς, ανεξάρτητος από οποιαδήποτε προηγούμενη γνώση ενός μονοδιάστατου μετρήσιμου συνόλου άκρου Χάουσοντορφ. Στο μεταξύ, το (3.72) θα μπορούσε να θεωρηθεί επίσης σαν ένας αλγόριθμος για εξαγωγή της στήριξης μιας μοναδικής μέτρησης  $D_1 u$ .

Το αξιοσημείωτο γεγονός σχετικά με τις εικόνες  $BV$  είναι ότι το σύνολο άλματος  $S_u$  μεταφέρει πράγματι μια καλή ιδιότητα κανονικότητας. Για την ακρίβεια, μετρώντας ένα μηδενικό σύνολο  $\mathcal{H}^1$ , το  $S_u$  είναι μια μετρήσιμη ένωση επανορθώσιμων καμπύλων [137]:

$$S_u \setminus (\text{an } \mathcal{H}^1 \text{ null set}) = \bigcup_{k=1}^{\infty} \gamma_k,$$

όπου κάθε  $\gamma_k$  είναι ένα συμπαγές υποσύνολο μιας καμπύλης  $C^1$ . Ως αποτέλεσμα, κάποιος θα μπορούσε να ορίσει το άλμα  $[u] = u_+ - u_-$  και ένας κατάλληλος κανονικός φορέας πεδίου  $\nu$  κατά μήκος του  $S_u$  μετρώντας ένα μηδενικό σύνολο  $\mathcal{H}^1$ . Τότε θα μπορούσε να αποδειχθεί ότι

$$D_1 u = [u] \nu \, d\mathcal{H}^1 = J_u,$$

στηριζόμενο στο  $S_u$ .

Συνοπτικά, έχουμε δείξει ότι ο φορέας μέτρηση Ρέιτον  $Du$  μιας γενικής εικόνας  $BV \mathbf{u}$  μπορεί να εκφραστεί ως

$$Du = \nabla u + J_u + D_c u.$$

Σε σύγκριση με τη φόρμουλα αποσύνθεσης (3.70) για μια ιδανική εικόνα Μάμφορτ-Σια, εκτός από τη λεπτή διαφορά ανάμεσα στο κλειστό μονοδιάστατο σύνολο άκρου  $\Gamma$  Χάουστορφ και το σύνολο άλματος  $S_u$ , μια εικόνα  $BV \mathbf{u}$  φαίνεται να είναι ένας ιδανικός υποψήφιος για το μοντέλο Μάμφορτ-Σια με την προϋπόθεση ότι ο συντελεστής Κατανομέα  $D_c u$  εξαφανίζεται. Αυτή είναι η παρατήρηση κλειδί που οδηγεί σε πολλά θεωρητικά αποτελέσματα σχετικά με το μοντέλο τμηματοποίησης Μάμφορτ-Σια.

**Ορισμός 3.27 (Ειδικές Εικόνες BV).** Μια εικόνα  $BV \mathbf{u}$  στο  $\Omega$  λέγεται ότι είναι ειδική αν ο συντελεστής Κατανομέα εξαφανίζεται παντού. Το υποδιάστημα όλων αυτών των ειδικών εικόνων  $BV$  υποδηλώνεται από το

$$SBV(\Omega).$$

υποδηλώνεται από το

Χαλαρώνοντας το συμπαγές αθροισματικό- $\mathcal{H}^1$  σύνολο άκρου  $\Gamma$  στο σύνολο άλματος  $S_u$ , κάποιος λαμβάνει την αδύναμη ενέργεια Μάμφορτ-Σια στο  $SBV(\Omega)$ :

$$E[u | u \in SBV, u_0 \text{ δεδομένου}] = \alpha \mathcal{H}^1(S_u) + \beta \int_{\Omega} |\nabla u|^2 dx + \gamma \int_{\Omega} (u_0 - K[u])^2 dx. \quad (3.73)$$

Η ύπαρξη αυτής της αδύναμης τμηματοποίησης Μάμφορτ-Σια επιβεβαιώνεται στη δουλειά του Αμβρόσιο [7]. Με

παραπέρα χαρακτηρισμό των λεπτών ιδιοτήτων του συνόλου άλματος  $S_u$  της αδύναμης λύσης οι Ντε Τζιόρτζι, Καριέρο, και Λεάσι [136], Ντελ Μάσο, Μορέλ, και Σολιμίνι [94], και Μορέλ και Σολιμίνι [221] απέδειξαν ότι το  $(u, \Gamma = S_u)$  είναι στην πραγματικότητα ένας ελαχιστοποιητής για την αρχική δυναμική τυποποίηση. Το Κεφάλαιο 7 ερευνά περαιτέρω το πασίγνωστο μοντέλο των Μάμφορτ-Σια.

# ΚΕΦΑΛΑΙΟ 5

ΑΠΟΒΟΛΗ ΘΟΛΩΜΑΤΟΣ ΑΠΟ ΜΙΑ ΕΙΚΟΝΑ (*Image Deblurring*)

Το πρόβλημα της ανακτησης μιας ευκρινούς εικόνας από την θαμπη εκδοχή της, είναι γνωστό ως *image deblurring*. Όπως η αποκοπή «ηχού» από μια εικόνα (*image denoising*), έτσι και η αποκοπή της θαμπαδας από μια εικόνα (*image deblurring*) προκύπτει συχνά στις επιστήμες και τις τεχνολογίες απεικόνισης,

συμπεριλαμβανομένων των οπτικών, ιατρικών, και αστρονομικών εφαρμογών, και είναι συχνά ένα κρίσιμο βήμα προς την επιτυχή ανίχνευση των σημαντικών σχεδίων όπως οι ανώμαλοι ιστοί ή τα επιφανειακά στοιχεία μερικών απόμακρων πλανητών.

Από μαθηματική άποψη, η αποκοπή της θαμπαδας από μια εικόνα, συνδέεται στενά με τις αντιστροφές διαδικασίες διαχυσης (π.χ., αντιστροφή της εξίσωσης της θερμότητας), οι οποίες είναι εμφανώς ασταθείς. Όπως οι αντιστροφές λύσεις προβλημάτων, έτσι και τα μοντέλα αποβολής θαμπαδας εξαρτώνται σε κρίσιμο βαθμό από τις καταλληλές βελτιωτικές τεχνικές, οι οποίες θα βοηθήσουν στην εξασφάλιση της σταθερότητας, συχνά όμως με την αναποφευκτική απώλεια ορισμένων, υψηλής συχνότητας, λεπτομερειών των ιδανικών εικόνων. Τέτοιες τεχνικές βελτίωσης μπορούν να έχουν σαν αποτέλεσμα την υπαρξη ή την μοναδικότητα μη θολωμένων εικόνων (*deblurred images*).

Σε αυτή την εργασία θα παρουσιαστούν τα φυσικά θεμέλια ορισμένων κοινών τυπών θολομάτος, θα ταξινομηθούν τα προβλήματα που προκύπτουν από την αποβολή της θαμπαδας από μια εικόνα και θα γίνει μια μαθηματική ανάλυση τόσο των μοντέλων ξεθολώματος όσο και των συναφών υπολογιστικών τους μεθόδων.

## **5.1 ΘΟΛΩΜΑ ΕΙΚΟΝΑΣ: ΦΥΣΙΚΗ ΠΡΟΕΛΕΥΣΗ ΚΑΙ ΜΑΘΗΜΑΤΙΚΑ ΜΟΝΤΕΛΑ**

### **5.1.1 ΦΥΣΙΚΗ ΠΡΟΕΛΕΥΣΗ**

Υπάρχουν τρεις σημαντικές κατηγορίες θολωματος σύμφωνα με το φυσικό τους υποβαθρο: *οπτικό, μηχανικό, και μεσοπροθεςμης προκλησεως.*

Το *οπτικο θολωμα* αποκαλείται, επίσης, και *εκτος εστιασης θολωμα*, και οφείλεται στην απόκλιση του επιπέδου απεικόνισης από το επίκεντρο ενός οπτικού φακού. Για παραδειγμα, για ένα μυωπικο ματι, ο αμφιβληστροειδής πέφτει ελαφρώς πίσω από την εστίαση της κορης του οφθαλμου. Από την άλλη πλευρά, κατά τη λήψη μιας υπαίθριας σκηνης με πολλά αντικείμενα σε ιδιαίτερα διαφορετικές σειρές, ο φακός μιας ψηφιακής φωτογραφικής μηχανής μπορεί να επικεντρωθεί σε ένα μόνο στόχο ή μια ομάδα ενδιαφεροντος, αφήνοντας τα υπόλοιπα στοιχεία εκτός εστίασης.

Το *μηχανικο θολωμα* προκύπτει από τις σύντομες μηχανικές κινήσεις, είτε των αντικειμένων – στόχων, είτε των συσκευών απεικόνισης, κατά τις διεργασίες λήψης εικόνας. Και οι δύο τύποι κινήσεων θα μπορούσαν να οδηγήσουν στο *θολωμα κινησης*, όπως συχνά αποκαλείται στη βιβλιογραφία. Όταν η σχετική κίνηση μεταξύ μιας συσκευής απεικόνισης και των στοχοθετημένων αντικειμένων της είναι αρκετά έντονη, η εικόνα ενός σημείου – αντικειμένου θα μπορούσε να μεταδοσει ευρεως πολλά εικονοστοιχεια στο επιπεδο απεικονισης κατά τη διάρκεια μιας μονο έκθεσης. Η εν λογω διασπορα καθως και η μιξη των πληροφοριων του χωρου αποτελουν τα θεμελια για *μηχανικο* ή *κινητικο θολωμα* εικονας.

Το *μεσοπροθεςμης προκλησεως θολωμα*, οφείλεται στη διασπορά ή στην οπτική διαταραχή των φωτονιακών μέσων, μέσω των οποίων ταξιδευουν οι ακτίνες φωτος. Η επίδραση της ατμόσφαιρας στις δορυφορικες εικόνες είναι ένα καλά μελετημένο παράδειγμα και συχνά ονομάζεται *ατμοσφαιρικο θολωμα*.

Οι χωροχρονικες διακυμανσεις των φυσικων ιδιοτητων, όπως η θερμοκρασία και η πυκνότητα της ατμόσφαιρας μπορεί να οδηγήσει σε τυχαία κυμαινόμενες διανομές του δείκτη διάθλασης, ενα φαινόμενο γνωστό ως *οπτική αναταραχη*.

Εν τω μεταξύ, χημικές ουσίες, διαφορετικών μεγεθών, μπορούν επίσης να εισαγάγουν διάφορους τύπους διασπορών στα φωτεινά κύματα με διαφορετικά μήκη κύματος. Στις ατμοσφαιρικές επιστήμες, το φαινόμενο συχνά αποκαλείται διασπορά αερολύματος, δηλαδή, διασπορά η οποια οφείλεται σε διάφορα χημικα ειδη που

αιωρούνται στον αέρα. Γνωστά παραδείγματα, περιλαμβάνουν το αλάτι λίγο πάνω από το επίπεδο της θάλασσας, στην θαλάσσια ατμόσφαιρα, την σκόνη από το έδαφος, στο ηπειρωτικό κλίμα, καθώς και διάφορες χημικές ουσίες προερχόμενες από τη βλάστηση. Στο φως του ήλιου, για παράδειγμα, σύννεφα και ομίχλη εμφανίζονται συχνά λευκά, δεδομένου ότι περιέχουν κυρίως σωματίδια μεγάλου μεγέθους και τα περισσότερα κύματα ορατού φωτός είναι εξίσου διεσπαρμένα.

Τόσο η οπτική αναταραχή, όσο και η διασπορά αερολύματος μπορούν να οδηγήσουν στην μίξη των πληροφοριών του χώρου, ή τη θαμπάδα. Η ατμοσφαιρική θαμπάδα αποτελεί ένα σημαντικό θέμα στις ατμοσφαιρικές επιστήμες, τη δορυφορική απεικόνιση, και τους απομακρισμένους αισθητήρες.

### 5.1.2 ΜΑΘΗΜΑΤΙΚΑ ΜΟΝΤΕΛΑ

#### **ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΘΟΛΩΜΑΤΟΣ ΚΙΝΗΣΗΣ**

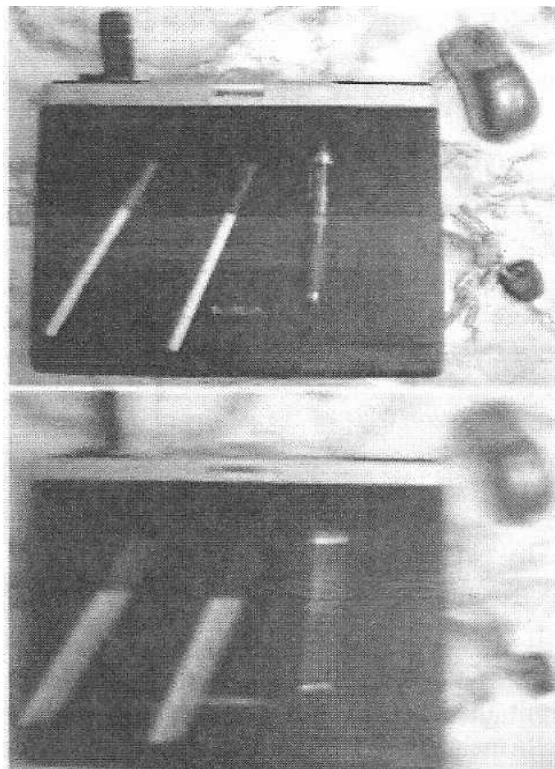
Μέσα σε μια πολύ σύντομη περίοδο από το χρόνο έκθεσης, οι περισσότερες κινήσεις μπορούν να προσεγγιστούν καλά από τις σταθερές ταχύτητες, όπως δικαιολογούνται από την επέκταση του Taylor. Πρέπει, όμως, να γίνει ακόμα η διακριση δύο διαφορετικών τύπων σταθερών κινήσεων, εκείνων των αντικειμένων στις τρισδιάστατες σκηνές και εκείνων των συσκευών απεικόνισης (π.χ., ψηφιακές κάμερες και βιντεοκάμερες). Ο τελευταίος τύπος προκαλείται συχνά από τις μηχανικές δονήσεις ή το ξαφνικό τρεμουλο των χεριών κατά τη διάρκεια μιας διαδικασίας λήψης εικόνας και οδηγεί στο ομοιόμορφο θόλωμα ολόκληρης της εικόνας. Η χωρική ομοιομορφία είναι χαρακτηριστική για αυτόν τον τύπο κινητικού θολώματος (βλεπε σχήμα 5.1).

Σύμφωνα με το πρώτο είδος της κίνησης, από την άλλη πλευρά, όταν οι συσκευές απεικόνισης διατηρούνται στατικές, οι προκληθείσες κινητικές θαμπάδες είναι χαρακτηριστικά ανομοιόμορφες, δεδομένου ότι τα διαφορετικά αντικείμενα μπορούν να κινηθούν σε ευδιάκριτες ταχύτητες. Για παράδειγμα, τα μακρινά υπόβαθρα είναι, συχνά, περίπου στατικά όσον αφορά μια ψηφιακή κάμερα, ενώ τα αυτοκίνητα τα οποία τρέχουν ή ένας παίχτης μπάσκετ ο οποίος κινείται, εντός μικρότερης

απόστασεως κινούνται αισθητά πιο γρήγορα. Κατά συνέπεια, τέτοιες κινητικές θαμπάδες είναι, χαρακτηριστικά, εντοπισμένες στις εγγύτητες των μεμονωμένων αντικειμένων.

Ενώ οι δύο τύποι των κινήσεων οδηγούν σε διαφορετικές σφαιρικές θολές μορφές, οι υποκείμενοι τοπικοί μηχανισμοί θολωματος είναι, όμως, ουσιαστικά οι ίδιοι. Αγνοώντας τη φυσική πολυπλοκότητα των φακών απεικόνισης, καθώς και τα στοιχεία κίνησης που είναι κάθετα προς τα επίπεδα απεικόνισης (τα οποία σχετίζονται με εκτός εστίασης θολωμα), τέτοια τοπικά θολώματα θα μπορούσαν να διαμορφωθούν ως εξής.

Ας υποθέσουμε ότι σε σκούρο φόντο, ένα μόνο φωτεινό σημείο (η εικόνα ενός φυσικού σημείου) μετακινείται από το αρχικό σημείο  $O$  σε ένα τελικό σημείο  $A$  κατά τη διάρκεια μιας μόνο έκθεσης σε κάποια ομοιόμορφη ταχύτητα. Έστω ότι  $[O, A]$  χαρακτηρίζει το τμήμα ευθειων γραμμων και  $L = |A - O|$  την απόσταση ταξιδιού στο επίπεδο απεικόνισης (βλεπε Σχήμα 5.2). Επιπλέον, το  $T$  υποδηλώνει την πάροδο του χρόνου μιας μόνο έκθεσης και το  $I$  την ένταση φωτισμού ενός φωτεινου φυσικου σημειου.



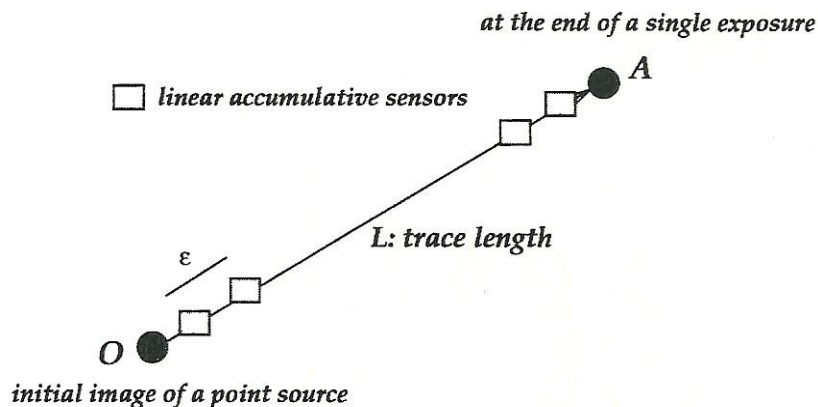
**Σχήμα 5.1.** Ένα παράδειγμα θολωματος κίνησης (λόγω τρεμουλου της καμερας).

Ας υποθέσουμε ότι κάθε αισθητήρας απεικόνισης είναι γραμμικός και συσσωρευτικός, έτσι ώστε χωρίς κίνηση η απάντηση της εικόνας σε  $O$  μετά από μία μόνο έκθεση θα είναι ανάλογο τόσο με το  $I$  και όσο και με το  $T$  :

$$u(O) = \mu \times I \times T = \mu \int_0^T I dt,$$

όπου  $\mu$  αποτελεί μια σταθερα απαντησης η οποία καθορίζεται από τη συσκευή απεικόνισης. Η υπόθεση ότι είναι συσσωρευτική είναι συχνά μια έγκυρη προσέγγιση κάτω από τις περισσότερους κανονικούς ορους σηματοδότησης ή συνθήκες φωτισμού. Για παράδειγμα, πριν από την πυροδότηση, το σώμα ενός νευρώνα συχνά μοντελοποιείται ως συσσωρευτικο στη μαθηματική βιολογία και στα νευρωνικά δίκτυα, δηλαδή, προσθέτοντας όλες τις απαντήσεις που έλαβε από δενδρίτες του.

Τώρα θεωρούμε καποια σταθερη κίνηση με ταχύτητα  $v = L / T$ . Σημειωνουμε όλους τους αισθητήρες απεικόνισης απο  $1, \dots, N$  κατά μήκος  $[O, A]$ . Ας υποθέσουμε ότι κάθε αισθητήρας καλύπτει μια αποτελεσματική φυσική  $\epsilon$  απόστασης από το επίπεδο απεικόνισης και  $L = N \times \epsilon$  (Σχήμα 5.2). Στη συνέχεια η απάντηση της εικόνας kn όπου ο αισθητηρας θα είναι





**Σχημα 5.2.** Θολωμα κινήσης μιας εικόνας από μια σημειακή πηγή

Ειδικότερα,

$$u(O) = \sum_{n=1}^N k_n, \quad (5.1)$$

το οποίο είναι το μοντέλο θολώματος για ένα ενιαίο σημείο - μια κατά τα άλλα στατική πηγή σημείου, η οποία διαδίδεται τώρα έξω και κατά μήκος του ίχνους της.

Στο όριο συνέχειας, το παραπάνω μοντέλο σημείων υποδηλώνει ότι ένα σημείων μονάδας προέλευσεως  $O$ , όπως παρουσιάστηκε από τη μέτρηση δέλτα του Dirac, όπου  $\delta(x) = \delta(x, y)$  είναι απλωμένο στο ομοιόμορφο μέτρο  $k(x)$  κατά μήκος του διαστήματος το  $[O, A]$ . Σημειώνουμε ότι εάν το  $[O, A]$  περιστρέφεται στο πρότυπο οριζόντιο διάστημα  $[0, L] \times \{0\}$  κατά μήκος του άξονα  $x$ , το ενιαίο μέτρο μπορεί να εκφράζεται με

$$k(x) = \frac{1}{L} 1_{[0,L]}(x) \times \delta(y).$$

Γενικά, θεωρούμε ότι το  $t$  χαρακτηρίζει το μοναδιαίο διάνυσμα της ταχύτητας της κίνησης και  $n$  τη μονάδα της κανονικής καθέτου στο  $t$ . Στη συνέχεια, το μοναδιαίο ποσό των πληροφοριών στο  $O$ , δηλαδή,  $\delta(x - O)$ , εκτείνεται στο γραμμικό μετρο

$$k_{O,t}(x) = \frac{1}{L} 1_{[0,L]}((x - O) \cdot t) \times \delta((x - O) \cdot n) = k_t(x - O), \quad (5.2)$$

όπου  $k_t(x)$  ανταποκρίνεται στην διασπορα του  $O = (0,0)$ .

Για την επέκταση ενός τέτοιου ενιαίου μοντέλου σημείων θολώματος στο σύνολο του επιπέδου απεικόνισης, αρχικά υποθέτουμε ότι κατά τη διάρκεια μίας μόνο έκθεσης ολόκληρη η σκηνή της εικόνας κινείται σε μια σταθερή ταχύτητα, με  $t$  κατεύθυνση και με απόσταση  $d$ . Ας υποθέσουμε ότι η ιδανική στατική εικόνα είναι  $u(y)$ , η οποία ορίζεται σε ολόκληρο το επίπεδο απεικόνισης του  $\mathbb{R}^2$ . Θεωρούμε κάθε pixel στόχου  $x$  της εικόνας σαν έναν αισθητήρα απεικόνισης. Η γραμμική υπόθεση για τους αισθητήρες υποδηλώνει ότι τη συμβολή μιας  $y$  pixel πηγης έως το pixel πεδίο  $x$  αποτελεί το ποσό του

$$u(\mathbf{y}) \times k_{\mathbf{y},t}(\mathbf{x}) = u(\mathbf{y}) \times k_t(\mathbf{x} - \mathbf{y}),$$

συμφωνα με το (5.2). Υποθετουμε ότι το  $u_0(\mathbf{x})$  χαρακτηριζει την θολη εικονα σε ένα σημειακο πεδιο  $x$ . Τοτε

η συνέλιξη της ιδανικής στατικής  $u$  εικόνας με το PSF (σημείο διάδοσης λειτουργίας) του πυρήνα

$$k_t(\mathbf{x}) = \frac{1}{L} 1_{[0,L]}(\mathbf{x} \cdot \mathbf{t}) \times \delta(\mathbf{x} \cdot \mathbf{n}).$$

Είναι, επομένως, προφανές ότι η (5,3) ισοδυναμει με το γραμμικο ολοκληρωμα

$$u_0(\mathbf{x}) = \frac{1}{L} \int_0^L u(\mathbf{x} - s\mathbf{t}) ds = \int_{\mathbb{R}^2} u(\mathbf{y}) d\mu_{\mathbf{x},t}(\mathbf{y}), \quad (5.4)$$

όπου το  $d\mu_{\mathbf{x},t}$  δηλώνει την ενιαία γραμμη μετρησης στο διάστημα  $[x - Lt, x]$ .

Στην συνεχεια, κάνουμε δύο ακόμη παρατηρήσεις πριν από την ολοκλήρωση της συζήτησης σχετικά με το μοντελο θολωματος κίνησης.

1. Έστω οτι οι αισθητήρες απεικόνισης είναι γραμμικοι, αλλά η κίνηση κατά μία μόνο έκθεση δεν είναι ομοιόμορφη σε ολόκληρο το πλάνο απεικόνισης. Στη συνέχεια η  $t$  κατεύθυνση και η συνολική απόσταση μεταφορας  $d$ , θα μπορούσαν εξαρτηθουν από τα μεμονωμένα pixel, και κατά συνέπεια, τα ενιαία μέτρα μετρησης  $d\mu_{\mathbf{x},t}$  στο μοντελο θολωματος (5.4) δεν είναι πλέον αναλλοίωτα. Αυτή είναι η περίπτωση, όπου το θολωμα κίνησης δημιουργειται κυρίως από μεμονωμένα αντικείμενα, σε γρήγορες κινήσεις.

2. Στην πραγματικότητα, οι αισθητήρες απεικόνισης, μπορεί συχνά να μην είναι γραμμικοι. Για παράδειγμα, οι φωτοϋποδοχείς του ανθρώπινου αμφιβληστροειδους, είναι γνωστό οτι είναι λογαριθμικοι, λόγω του νόμου του

Weber. Επιπλέον, πολλοί ψηφιακοί ή βιολογικοί αισθητήρες απεικόνισης, είναι συνήθως κορεσμένοι σε ορισμένα κατώτερα επίπεδα, οδηγώντας σε ένα άλλο είδος μη γραμμικής πολυπλοκότητας. Για παράδειγμα, ας υποθέσουμε ότι η ένταση φωτισμού  $I$  ενός μόνο φωτεινού σημείου, είναι τόσο ισχυρή που το  $\varepsilon/L \times I$  υπερβαίνει ήδη το επίπεδο κορεσμού. Τότε, ο τύπος διαδοσης (5.1), αναθεωρείται ώστε

$$u(O) = k_1 = \dots = k_N \equiv \text{η κορεσμένη αντίδραση των αισθητήρων,}$$

το οποίο δεν είναι μέσου όρου ή χαμηλής διέλευσης, και, κατά συνέπεια, δεν μπορεί να εκφραστεί από την συνέλιξη. Στην περίπτωση αυτή, ο στόχος φωτεινό σημείο απλώς αντιγράφεται στο δρόμο του.

#### ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΕΚΤΟΣ ΕΣΤΙΑΣΗΣ ΘΟΛΩΜΑΤΟΣ

Ας εξετάσουμε το ενδεχόμενο ενός διαφράγματος απεικόνισης μέσω ενός, χωρίς παρεκκλιση, και κυρτού λεπτού φακού. Έστω ότι το  $d$  χαρακτηρίζει την απόσταση από ένα σημείο αντικείμενο στο πλάνο του φακού, το  $f$  την απόσταση εστίασης του φακού, και το  $e$  την απόσταση από το σημείο της εικόνας στο το πλάνο του φακού (βλ. σχήμα 5.4). Τότε ο ιδανικός νομος για τον φακο αναφέρει ότι

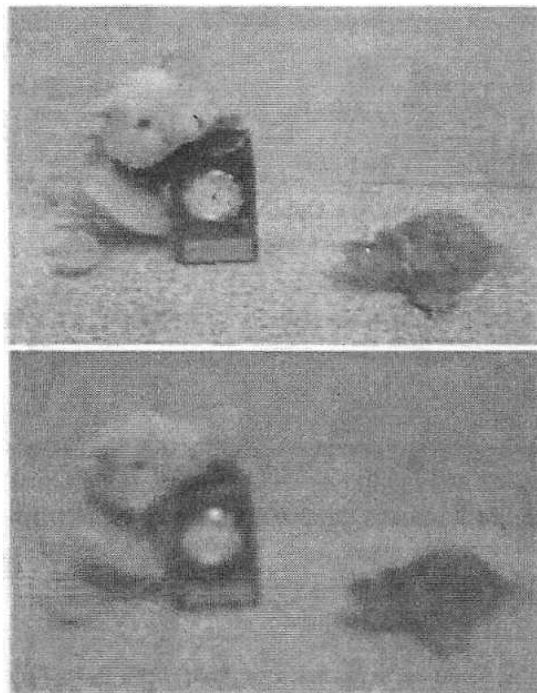
$$\frac{1}{d} + \frac{1}{e} = \frac{1}{f}. \quad (5.5)$$

Έστω ότι το  $h$  χαρακτηρίζει την απόσταση μεταξύ του επιπέδου του φακού και του επιπέδου απεικόνισης. Εάν  $h = e$ , τότε η εικόνα του σημείου αντικειμένου είναι σαφώς εστιασμένη, στο επίπεδο απεικόνισης. Σε αντίθετη περίπτωση, ένα ενιαίο φυσικό σημείο σε απόσταση  $d$ , θα ήταν θολο, σε ένα μεγάλο σημείο, στο επίπεδο απεικόνισης (Σχήμα 5.4).

Ομοια με τα τρίγωνα, είναι εύκολο να δείξουμε στο σχήμα 5.4 ότι αυτό το θολο σημείο, είναι μία περιοχή όσο το χείλος του φακού, το οποίο είναι κυκλικό. Έστω ότι

το  $r$  και το  $a$  δηλώνουν τις ακτίνες του θολου σημείου και του φακού χωριστά. Τότε, θα περίμενε κανείς ότι μεταξύ των παραμέτρων υπάρχει κάποια σχέση, όπως η

$$r = r(a, f, d, h).$$



**Σχήμα 5.3.** Παραδειγμα εκτος εστιασης θλωματος

Στην πραγματικότητα, η ομοιότητα των δύο τριγώνων OAD και BCD στο σχήμα 5.4,

$$\frac{r}{a} = \frac{h - e}{e} = h \left( \frac{1}{e} - \frac{1}{h} \right).$$

Σύμφωνα με το νόμο του φακού (5.5), προκύπτει ένας σαφής τύπος

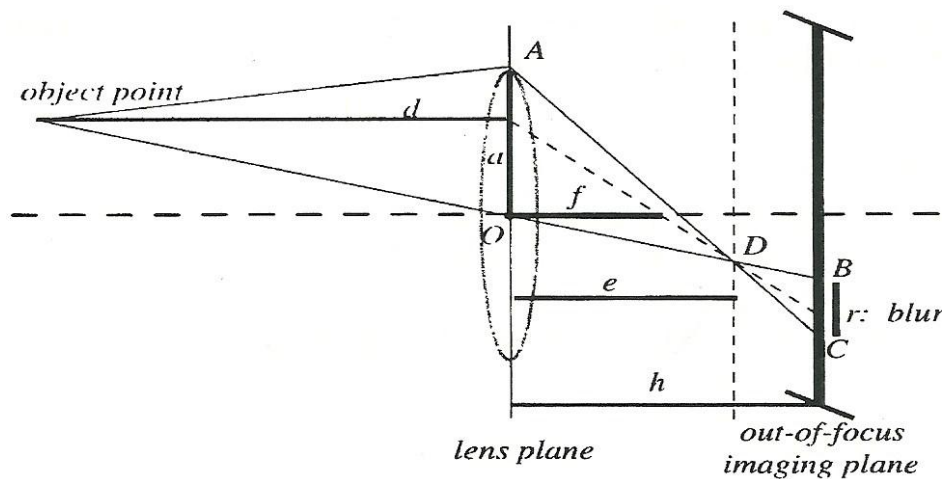
$$r = ah \left( \frac{1}{f} - \frac{1}{d} - \frac{1}{h} \right). \quad (5.6)$$

Κατά συνέπεια, η οξεία όραση εμφανίζεται χωρίς την θαμπαδα,  $r = 0$ , εάν και μόνον εάν, ο νομος των φακών υπάρχει. Για τις περισσότερες συσκευές απεικόνισης, οι σταθερές  $a$  και  $f$  των φακων καθορίζονται. Τότε ο βαθμός του θαμπώματος είναι συνάρτηση του  $d$  και του  $h$  μόνο.

Συνοψίζοντας, για ένα ορισμένο, εκτός εστιασης, επιπεδο απεικονησης και οποιαδήποτε αντικείμενο σε ένα σταθερό φασμα επιπεδου, το PSF αποτελεί τον δεικτη λειτουργιας ενος δίσκου με ακτίνα  $r$ :

$$k(\mathbf{x}) = k_r(\mathbf{x}) = \frac{1}{\pi r^2} 1_{B_r}(\mathbf{x}), \quad B_r = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| < r\}.$$

Στην πραγματικότητα, η ανωτέρω ανάλυση, βασισμένη και στη γεωμετρική οπτική και στην ιδανική υπόθεση των φακών, οφείλεται, μόνο προσεγγιστικως, στο φαινόμενο διάθλασης και την οπτική παρέκκλιση.



Σχημα 5.4. Γεωμετρική οπτική της εκτός εστιασης απεικονησης.

Τα διαφορετικά μήκη κύματος προξενούνται με διαφορετικούς ρυθμούς και μπορούν να ακυρώσουν ή να ενισχύσουν το ένα το άλλο, έτσι ώστε το καθαρό, εκτός εστίασης, αποτέλεσμα να είναι μια ομαλά αποσυντιθειμενη κυκλική άμορφη μάζα στο επίπεδο απεικόνισης, αντί μιας ευδιακριτης περιοχης δισκων που λαμβάνεται μέσω γεωμετρικής οπτικής.

Όπως έχει μελετηθεί σε πολλές εργασίες, η θολή άμορφη μάζα θα μπορούσε να προσεγγιστεί και από το Γκαουσιανό μοντελο (Gaussian PSF):

$$k_{\sigma}(\mathbf{x}) = k_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right).$$

Η στατιστική ακτίνα  $\sigma$  είναι ανάλογη με την γεωμετρική ακτίνα  $r$  του (5,6):  $\sigma = \rho r$ , με την σταθερά  $\rho$  να καθορίζεται από τη συσκευή απεικόνισης. Έτσι

$$\sigma = \rho ah \left( \frac{1}{f} - \frac{1}{d} - \frac{1}{h} \right),$$

καθώς και κάθε άλλη σαφώς εστιασμενη εικόνα  $u(\mathbf{x})$  είναι θολή για

$$u_0(\mathbf{x}) = k_{\sigma} * u(\mathbf{x}) = \int_{\mathbb{R}^2} u(\mathbf{y}) k_{\sigma}(\mathbf{x} - \mathbf{y}) d\mathbf{y} \quad (5.7)$$

στην ιδανική ρύθμιση, όταν το επίπεδο απεικόνισης είναι  $\mathbb{R}^2$ . Αυτό είναι το πιο συχνά χρησιμοποιουμενο μοντέλο θολωματος, στην επεξεργασία εικόνας. Το σχήμα 5.3 δείχνει ένα πραγματικό παράδειγμα του εκτός εστίασης θολωματος.

## ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΤΟΥ ΑΤΜΟΣΦΑΙΡΙΚΟΥ ΘΟΛΩΜΑΤΟΣ

Λόγω της πολυπλοκότητας των ατμοσφαιρικων αναταράξεων και της διάχυση των , αερολυμάτων, η μαθηματική μοντελοποίηση του ατμοσφαιρικου θολωματος έχει πολύ μεγαλύτερη συμμετοχή από τις δύο προηγούμενες περιπτώσεις, και έχει μελετηθεί με προσοχή σε πολλούς σημαντικούς τομείς, συμπεριλαμβανομένων των

προσαρμοσιμων οπτικών συστημάτων, τις δορυφορικές εικόνες, και τους απομακρυσμένους αισθητήρες.

Στα περισσότερα μοντέλα, η ατμόσφαιρα προσεγγίζεται με ένα γραμμικό σύστημα, και το ατμοσφαιρικό θόλωμα μοντελοποιείται ώστε να είναι μετατόπιστικά αμετάβλητο, με ένα λειτουργικό σημείο εξάπλωσης (PSF)  $k(x)$ , έτσι ώστε μια θολή εικόνα  $u_0(x)$  που παρατηρήθηκε, να σχετίζεται με την ιδανική έντονη εικόνα  $u(x)$  από

$$u_0(x) = k(x) * u(x).$$

Στην οπτική, ο μετασχηματισμός Fourier  $K(\omega) = K(\omega_1, \omega_2)$  του PSF  $k(x)$ , συχνά αποκαλείται *οπτική συνάρτηση μεταφοράς* (OTF), και τα περισσότερα μοντέλα ατμοσφαιρικού θολώματος, ορίζονται με βάση τα OTF τους.

Για παράδειγμα, όπως επωφελήθηκε από τη θεωρία των αναταράξεων του Kolmogorov, το OTF θα μπορούσε να απεικονιστεί από

$$K(\omega) = \exp\left(-\beta \|\omega\|^{\frac{5}{3}}\right). \quad (5.8)$$

Υπενθυμίζουμε ότι η μέθοδος μετασχηματισμού του Fourier ορίζεται από

$$K(\omega) = \int_{\mathbb{R}^2} k(x) e^{-i\omega \cdot x} dx.$$

Σε μια πιο συγκεκριμένη ρύθμιση της απεικόνισης του τηλεσκοπίου μέσα από ένα λεπτό οριακό στρώμα της ατμόσφαιρας, το PSF έχει επίσης μοντελοποιηθεί με  $k(x) = |Ae^{i\phi}|^2$ , το οποίο δηλώνει το αντίστροφο του μετασχηματισμού του Fourier. (5.9)

Εδώ το  $A$  είναι η λειτουργία διαφράγματος του τηλεσκοπίου και  $\phi$  ο παράγοντας του σταδίου. Η κατανομή του σταδίου  $\phi$  κωδικοποιεί το καθαρό οπτικό αποτέλεσμα ενός λεπτού στρώματος της ταραχώδους ατμόσφαιρας και μπορεί να ληφθεί με τη μέτρηση της παραμόρφωσης του σταδίου, σε ένα επίπεδο μέτωπο κύματος, από ένα μακρινό αστέρι ή μια κατευθυνόμενη ακτίνα λέιζερ. Οι σχετικές OTF είναι συχνά

περιορισμένες, δεδομένου ότι η λειτουργία διάφραγματος είναι συμπαγώς υποστηρίζομενη στο πεδίο ανάλυσης του Fourier.

### 5.1.3 ΓΡΑΜΜΙΚΑ ΕΝΑΝΤΙΟΝ ΜΗ ΓΡΑΜΜΙΚΩΝ ΘΟΛΩΜΑΤΩΝ

Όλα τα παραπάνω τρία μοντέλα είναι γραμμικά. Ένα γενικό γραμμικό θόλωμα έχει τη μορφή  $u_0 = K[u]$  με το  $K$  να είναι διαρκές για τον φορέα του γραμμικού θόλωματος. Στις περισσότερες εφαρμογές, ο θόρυβος είναι αναπόφευκτος, και μια πραγματική παρατήρηση συχνά δίνεται από

$$u_0 = K[u] + n,$$

υπό τον όρο ότι ο θόρυβος  $n$  είναι πρόσθετος.

Ενας γραμμικός φορέας  $K$  λέγεται ότι είναι μετατόπιστικά αμετάβλητος αν για κάποιο  $\mathbf{a} \in \mathbb{R}^2$ ,

$$u_0(\mathbf{x}) = K[u(\mathbf{x})] \quad \text{υποδηλώνει ότι} \quad u_0(\mathbf{x} - \mathbf{a}) = K[u(\mathbf{x} - \mathbf{a})].$$

Είναι γνωστό στην επεξεργασία των σημάτων, καθώς και στην θεωρία του συστήματος ότι ένας μετατόπιστικά αμετάβλητος γραμμικός φορέας πρέπει να έχει τη μορφή της συνέλιξης:

$$K[u] = k * u(\mathbf{x}) = \int_{\mathbb{R}^2} k(\mathbf{x} - \mathbf{y})u(\mathbf{y})d\mathbf{y} \quad (5.10)$$

για κάποια κατάλληλη λειτουργία  $k(\mathbf{x})$  του πυρήνα, η οποία αποκαλείται PSF στο πλαίσιο της ατμοσφαιρικής απεικόνισης. (Χωρίς επιπλέον πληροφορίες σχετικά με την κανονικότητα, ένας μετατοπιστικά αμετάβλητος γραμμικός φορέας, θα μπορούσε να είναι πολλοί διαφορετικοί φορείς, ή στο πεδίο ανάλυσης κατά τον Fourier, πολυωνυμικοί πολλαπλασιαστές.)

Γενικότερα, αν το  $K$  είναι γραμμικό, στη συνέχεια σε κάθε σταθερο εικονοστοιχείο  $\mathbf{x}$  στον τομέα εικονοστοιχείων  $\Omega$ ,



$$L_x : u \rightarrow K[u](x)$$

πρέπει να υπάρχει μια γραμμική λειτουργική για το  $u$ , ή μια γενικευμένη συνάρτηση. Υποδηλώνει συμβολικά με  $k(x, \cdot)$ , έτσι ώστε, όπως στη θεωρία κατανομής, το ένα έχει

$$L_x[u] = \langle k(x, \cdot), u(\cdot) \rangle.$$

Ας υποθέσουμε ότι η διανομή  $k(x, \cdot)$  στην πραγματικότητα μπορεί να αναπαρασταθεί με μια συνηθισμένη μετρήσιμη λειτουργία στο  $L^1(\Omega)$ . Στη συνέχεια το γραμμικό θόλωμα μετατρέπεται σε θέση εξαρτώμενη από το συνηθισμένο ολοκλήρωμα:

$$u_0(x) = \int_{\Omega} k(x, y)u(y)dy.$$

Για τους σκοπούς της ανάλυσης, το  $u$  μπορεί να θεωρηθεί ότι ανήκει στο  $L^p(\Omega)$  με  $p \in [1, +\infty]$ , και το  $K$  είναι ένας (οριοθετημένος) γραμμικός φορέας από  $L^p(\Omega)$  σε  $L^q(\Omega)$  με κάποιο  $q \in [1, +\infty]$ . Ως αποτέλεσμα, το συζυγές  $K^*$  ορίζεται από  $(L^q)^*$  σε  $(L^p)^*$ , τους διπλούς χώρους. (Κάποιος πρέπει να γνωρίζει, ωστόσο, ότι  $(L^\infty)^* \neq L^1$ . Στη συνέχεια, οι τιμές της  $p$  και  $q$  πρέπει να προσαρμόζεται όπως απαιτείται.

Οποιοσδήποτε γραμμικός φορέας δεν θα μπορούσε να ονομαστεί θόλωμα, είτε πρόκειται για μετατόπιστικά αμετάβλητο, είτε όχι. Η λιγότερο απαιτητική, για την ύπαρξη ενός θόλωματος του φορα, είναι η κατάσταση *συνεχούς ρευματος (DC)*:

$$K[1] = 1, \quad 1 \in L^\infty(\Omega). \quad (5.11)$$

Όπως και στην ηλεκτρική εφαρμοσμένη μηχανική, η κατάσταση DC ξεχωρίζει για το συνεχές ρεύμα δεδομένου ότι ο μετασχηματισμός Fourier μιας σταθεράς, δεν περιέχει συχνότητες διαφορετικές από το μηδέν. Με το διπλό τύπο για την συζυγία

$$\langle K[u], v \rangle = \langle u, K^*[v] \rangle,$$

η κατάσταση συνεχούς ρεύματος (DC) στην  $K$ , υπολογίζεται από την κατάσταση της μεσης διατήρησης στην  $K^*$ :

$$\langle K^*[v] \rangle = \langle v \rangle, \quad \text{by setting } u = 1; \quad \text{or} \quad \int_{\Omega} K^*[v](\mathbf{x})d\mathbf{x} = \int_{\Omega} v(\mathbf{x})d\mathbf{x}, \quad (5.12)$$

εάν τόσο το  $v$  όσο και  $K^*[v]$  ανήκουν στο  $L^1(\Omega)$ .

Με την κατάσταση συνεχούς ρεύματος, οι απλοί παράγοντες διαφοροποίησης, ιδίως, δεν μπορεί να αντιμετωπίζονται ως θαμπάδες, οι οποίες περιλαμβάνουν την κλίση  $\nabla$ , την απόκλιση  $\nabla \cdot$  και το Laplacian  $\Delta$ .

Για μια αμετάβλητη μετατοπιστικά γραμμική θόλωση, με  $k$  πυρήνα, η κατάσταση συνεχούς ρεύματος απαιτεί

$$\int_{\mathbb{R}^2} k(\mathbf{x})d\mathbf{x} = 1, \quad \text{ή υπο τους ορους του μετασχηματισμου Fourier, } K(0) = 1.$$

Επιπλέον, ένας πιο ουσιαστικός παραγοντας θόλωσης πρέπει να είναι χαμηλής διέλευσης, το οποίο σημαίνει ότι στην μετατοπιστικά αμετάβλητη περίπτωση το  $K(\omega)$  διασπάται γρήγορα για μεγάλες συχνότητες.

Τέλος, οι θαμπάδες δεν χρειάζεται να είναι γραμμικές, αν και τα γραμμικά μοντέλα είναι κυρίαρχα στη βιβλιογραφία. Σκεφτείτε, για παράδειγμα, το μοντέλο ελάχιστης επιφανειακής διάχυσης:

$$v_t = \nabla \cdot \left[ \frac{1}{\sqrt{1 + |\nabla v|^2}} \nabla v \right], \quad v|_{t=0} = u(\mathbf{x}). \quad (5.13)$$

Έστω ότι η λύση συμβολίζεται με  $v(\mathbf{x}, t)$ . Για οποιοδήποτε σταθερό πεπερασμένο χρόνο  $T > 0$ , ένας μη γραμμικός παραγοντας  $K$  καθορίζεται από

$$u_0 = K[u] = v(\mathbf{x}, T).$$

Η μη γραμμικότητα είναι εμφανής, δεδομένου ότι, για παράδειγμα,  $K[\lambda u] \neq \lambda K[u]$  για το γενικό  $u$  και  $\lambda \neq 0$ . Όμως, ο παραγοντας  $K$  ικανοποιεί προφανώς τον όρο συνεχούς ρεύματος. Επιπλέον, η (5.13) είναι η εξίσωση κλίσης καθόδου της ελάχιστης επιφανειακής ενέργειας

Έτσι, είναι προφανές ότι τα μικρής κλίμακας χαρακτηριστικά και οι ταλαντώσεις του  $u$  εξαλείφθηκαν στο  $\mathcal{U}_0 = K[\mathbf{u}]$ , κάνοντας το  $\mathcal{U}_0$  μια οπτικά θολή και μαλακωμένη έκδοση της αρχικής ιδανικής εικόνας  $u$ . Αυτά τα δεδομένα που εξαρτώνται από το μοντέλο μη γραμμικής θόλωσης είναι, ωστόσο, μετατοπιστικά αμετάβλητα.

## 5.2 ΚΑΚΗ ΤΟΠΟΘΕΤΗΣΗ ΚΑΙ ΤΑΚΤΟΠΟΙΗΣΗ

Η κακή τοποθέτηση του πρόβληματος ξεθολωματος, μπορεί να κατανοηθεί επαρκώς, από τέσσερις διακριτές αλλά στενά συνδεδεμένες, οπτικές γωνίες.

- ***Η διαδικασία ξεθολωματος πρόκειται να αναστρέψει το χαμηλής διέλευσης φίλτράρισμα***

Από την κλασική αποψη επεξεργασίας σήματος, στη φασματική περιοχή Fourier, ένας παραγοντας θολωματος είναι συχνά ένα χαμηλής διέλευσης φίλτρο που μειώνει τις, υψηλής συχνότητας, λεπτομέρειες, μέσω πολλαπλασιαστών εξαφάνισης. Ως εκ τούτου, για να ξεθολώσουμε μια θολή εικόνα, κάποιος πρέπει να πολλαπλασιάσει σε ορισμένο βαθμό τα αντίστροφα των πολλαπλασιαστών εξαφάνισης, τα οποία γίνονται εύλογα ασταθή σε θορύβους ή άλλες υψηλής συχνότητας μικρές διαταραχές στα δεδομένα εικόνας.

- ***Η αποβολή θολωματος είναι μια αντιστροφή διαδικασία διάχυσης***

Με την κανονική θεωρία των γραμμικών παραβολικών PDEs, για να θολωθεί μια εικόνα με ποσοτητες του Γκαουσιανού πυρήνα, λειτουργεί η εξίσωση διάχυσης θερμότητας για κάποιο χρονικό διάστημα με μια συγκεκριμένη εικόνα, όπως τα αρχικά δεδομένα. Ως εκ τούτου, το ξεθολωμα φυσιολογικά ισοδυναμεί με την αναστροφή της διαδικασίας διάχυσης, η οποία είναι γνωστή για την ασταθεια της.

- ***Το ξεθολωμα μειώνει την εντροπία***

Ο στόχος του ξεθλωματος είναι να ανακατασκευάσει τα λεπτομερή χαρακτηριστικά της εικόνας από μια μαλακωμενη θολή εικόνα. Έτσι, από τη στατιστική μηχανική άποψη, το ξεθλωμα είναι μια διαδικασία για την αύξηση της πληροφόρησης, ή ισοδύναμα, της μείωσης της εντροπίας, η οποία σύμφωνα με τον Δεύτερο Νόμο της στατιστικής μηχανικής δεν μπορεί ποτέ να συμβεί φυσικά. Δηλαδή, επιπλέον εργασία πρέπει να γίνει στο σύστημα, μιλώντας μηχανικά.

- ***Το ξεθλωμα αναστρέφει τους Συμπαγεις Φορείς***

Από την άποψη της αφηρημένης λειτουργικής ανάλυσης, μια διαδικασία θλωματος είναι συχνά ένας συμπαγης φορέας. Ένας συμπαγης φορέας είναι αυτός που χάρτογραφει οποιαδήποτε οριοθετούμενο σύνολο σε ένα πολύ καλύτερο διαχειριζόμενο σύνολο το οποίο είναι μη συμπαγες. Για την επίτευξη αυτού του στόχου, ένας συμπαγης φορέας πρέπει να αναμειγνύει χωρικές πληροφορίες ή να εισάγει πολλές συνεκτικές δομές, οι οποίες συχνά πραγματοποιούνται από ουσιαστική μείωση των διαστάσεων με βάση την εξαφάνιση των ιδιοτιμών ή των μοναδικων τιμών. Ως εκ τούτου, η αντιστροφή ενός συμπαγους φορέα είναι και πάλι ισοδύναμη με την μη συσχέτιση της χωρικής συνοχής ή την ανακατασκευή πρώην καταστέλλομενων διαστάσεων (κατά τη διάρκεια της διαδικασίας θλώματος) των χαρακτηριστικών και των πληροφοριών, τα οποία είναι ασταθή.

- ***Η συστηματοποίηση είναι ζωτικής σημασίας για το ξεθλωμα***

Εν κατακλείδι, κατάλληλες τεχνικές συστηματοποίησης πρέπει να εφαρμοστούν για την καλύτερη κατάσταση της διαδικασίας ξεθλωματος. Δύο καθολικές προσεγγίσεις συστηματοποίησης, οι οποίες ουσιαστικά είναι αμοιβαίες καθρεπτικές εικόνες στους δύο διπλούς κόσμους των προσδιοριστικών και στοχαστικών μεθόδων, είναι η Tikhonov συστηματοποίηση και η Μπεϋζιανη θεωρία της λήψης αποφάσεων ή εξαγωγής συμπερασμάτων. Η στενή σύνδεσή τους εξηγείται, για παράδειγμα, απο τους Mumford και Chan, Shen, και Vese. Στην ουσία, οι δύο προσεγγίσεις εισάγουν

κάποια εκ των προτέρων γνώση για τις εικόνες στόχους  $u$ , ώστε να ανακατασκευαστούν. Στο Μπεϋζιανό πλαίσιο, πρέπει να εισαχθεί κάποια σωστή κατανομή πιθανοτήτων για κάθε δυνατή υποψηφια εικόνα, και ως εκ τούτου η αναγκαία μεροληψία ενθαρρύνεται ώστε να προωθήθουν περισσότερες πιθανότητες. Στη ρύθμιση Tikhonov, η εκ των προτέρων γνώση αντανακλάται συχνά μέσα από κάποιες σωστά σχεδιασμένες "ενέργειακες" συνθέσεις.

### 5.3 ΞΕΘΩΛΩΜΑ ΜΕ ΦΙΛΤΡΑ WIENER

Στην ενότητα αυτή, ερευνούμε μοντέλα ξεθολωματος που βασίζονται σε φίλτρα Wiener. Σκεφτείτε μια γραμμική μετατόπιση - αμετάβλητο μοντέλο θολωματος με προσθετικό θόρυβο

$$u_0 = k * u + n,$$

όπου ο PSF  $k(x)$  θεωρείται ίσος με τον γνωστό μετασχηματισμό Fourier (ή OTF)  $K(\omega)$ . Από τώρα και στο εξής, αντί για τα μαύρισμενα εντονα τα σύμβολα  $x$  και  $\omega$ , ένα γενικό εικονοστοιχείο, θα συμβολίζεται με  $x = (x_1, x_2)$  και η διπλή του συχνότητα, με  $\omega = (\omega_1, \omega_2)$ .

#### 5.3.1 Ξεθολωμα βασιζομενο σε Φιλτρο

Το ξεθολωμα μέσω φιλτραρίσματος, αξιολογει την ιδανική σαφή εικόνα  $u$  από την θολή εκδοση της  $u_0$ , μεσω κάποιου φίλτρου εκτιμητή

$$\hat{u} = \hat{u}_w = w * u_0$$

με ένα καταλληλο φιλτρο  $w$ .

Για καθαρά απομάκρυνση θορύβου χωρίς θόλωμα, το γραμμικό φιλτράρισμα φαίνεται ένα φυσικό εργαλείο για την καταστολή του θορύβου μέσω κινητού μέσου

όρου (δηλαδή, την συνέλιξη). Για ξεθολωμα, ωστόσο, φαίνεται λίγο ριψοκίνδυνη προσπάθεια να αφαιρέσετε μια συνέλιξη (π.χ.,  $k^*$ ) από μια άλλη συνέλιξη (δηλαδή,  $w^*$ ). Για παράδειγμα στην χωρίς θόρυβο περίπτωση, η ιδανική λύση θα ήταν, όπως βολικά εκφράζεται στο πεδίο ανάλυσης κατά Fourier,

$$W(\omega) = \frac{1}{K(\omega)}, \quad (5.14)$$

έτσι ώστε  $\hat{u} = w^* u_0 \equiv u$  για οποιαδήποτε σαφή εικόνα  $u$ , και η τέλεια ανασυγκρότηση φαίνεται να είναι φθηνά καθορισμένη. Ωστόσο, η (5.14) είναι μια μάλλον ανησυχητική φόρμουλα, αφού ένα τυπικό  $k$  θόλωμα είναι συχνά ένα χαμηλής διέλευσης φίλτρο και το  $K(\omega)$  διασπάται γρήγορα σε υψηλές συχνότητες οδηγώντας σε ασταθή blowups στο φίλτρο ξεθολωματος  $W(\omega)$ . Ένα τέτοιο φίλτρο ξεθολωματος, ως εκ τούτου, μπορεί να ενισχύσει οποιοδήποτε υπερβολικά μικρό λάθος σε υψηλές συχνότητες. Για να ξεπεραστεί αυτός ο ενδεχόμενος κίνδυνος της ασταθούς αποκατάστασης, στην περίπτωση χωρίς θόρυβο, αναδιατυπώνουμε την (5.14) ως εξής

$$W(\omega) = \frac{K^*(\omega)}{K(\omega)K^*(\omega)} = \frac{K^*}{|K|^2}, \quad \text{όπου το } * \text{ δηλώνει συζυγή μιγαδικό.}$$

Στη συνέχεια, μπορεί κανείς να προσπαθήσει να διευθετήσει την εν δυνάμει εξαφάνιση παρονομαστή στις υψηλές συχνότητες με την προσθήκη ορισμένου θετικού παράγοντα  $r = r(\omega)$ :

$$W \rightarrow W_r = \frac{K^*}{|K|^2 + r}.$$

Έστω ότι το  $\hat{u}_0$  χαρακτηρίζει τον εκτιμητή αποτέλεσμα. Στη συνέχεια (στην χωρίς θόρυβο ρύθμιση)

$$\hat{u}_r = w_r * k * u,$$

ή στον τομέα ανάλυσης κατά Fourier, ο καθαρός φορέας ανασυγκρότησης γίνεται ο πολλαπλασιαστής

$$W_r(\omega)K(\omega) = \frac{|K(\omega)|^2}{|K(\omega)|^2 + r(\omega)}. \quad (5.15)$$

Με αυτόν τον τύπο, η αποκατάσταση πραγματι προσεγγίζει τον φορέα ταυτότητοποιησης σε χαμηλές συχνότητες, όπου  $r \ll |K|^2$ . Οι υψηλές συχνότητες, όμως, διαστρεβλώνονται από την καταστολή, δεδομένου ότι το  $K$  σχεδόν εξαφανίζεται και  $|K|^2 \ll r$ . Έτσι, το ρυθμιστικό  $r$ , διαδραματίζει έναν ρόλο οριακού ή κατώτατου όριου. Η εξίσωση (5.15) είναι ένα χαρακτηριστικό παράδειγμα για ρυθμισμένο αντίστροφο φίλτράρισμα. Τώρα το ερώτημα είναι πώς να επιλέξουμε συνετά το καλύτερο ρυθμιστικό  $r$ . Αποδεικνύεται ότι για το σκοπό αυτό ο θόρυβος μπορεί να συμβάλει σημαντικά αντί να είναι ενοχλητικός. Αυτό επιτυγχάνεται με την αρχή του Wiener, με το λιγότερο μέσο τετραγωνικό σφάλμα.

### 5.3.2 Φιλτραρισμα κατά Wiener

Το φίλτράρισμα κατα Wiener, υποθέτει ότι ο εκτιμητής ξεθολωματος  $\hat{u}$ , είναι μια φιλτραρισμένη εκδοχή της παρατήρησης  $u_0$ , χρησιμοποιώντας κάποιο βέλτιστο φίλτρο  $w(x)$ :

$$\hat{u}_w = w * u_0.$$

Το φίλτρο του Wiener  $w$ , είναι η ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος εκτίμησης  $e_w(x) = \hat{u}_w(x) - u(x)$ :

$$w = \underset{h}{\operatorname{argmin}} E[e_h^2] = \underset{h}{\operatorname{argmin}} E[(h * u_0(x) - u(x))^2]. \quad (5.16)$$

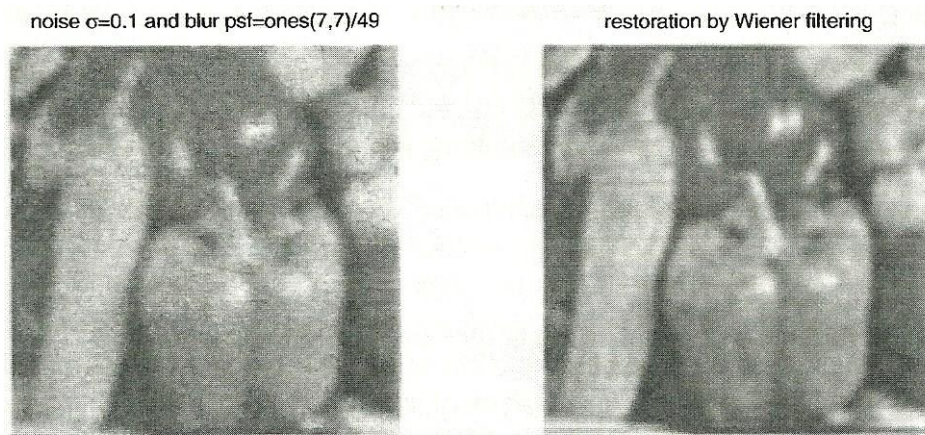
Όπως προκύπτει, η (5.16) οδηγεί στην ορθογώνια κατάσταση

$$E[(w * u_0(x) - u(x))u_0(y)] = 0 \quad \forall x, y \in \Omega, \quad (5.17)$$

η οποία από την άποψη των λειτουργιών συσχέτισης γίνεται

$$w * R_{u_0 u_0}(z) = R_{u u_0}(z), \quad z \in \mathbb{R}^2.$$





**Σχημα 5.5.** Ένα παραδειγμα του φίλτρου του Wiener τόσο για την αποκοπη του θορυβου όσο και για το ξεθολωμα.

Κατά συνέπεια, ο ρητος τύπος για το βέλτιστο φίλτρο του Wiener είναι

$$W(\omega) = \frac{S_{uu_0}(\omega)}{S_{u_0u_0}(\omega)}, \quad (5.18)$$

Για το μοντέλο θόλωματος  $u_0 = k * u + n$ , κάποιος έχει, σύμφωνα με θεώρημα 4.2,

$$S_{uu_0} = K^*(\omega)S_{uu}(\omega) \text{ and } S_{u_0u_0} = |K(\omega)|^2 S_{uu}(\omega) + S_{nn}(\omega).$$

Ως εκ τούτου, έχουμε αναπτύξει το παρακάτω θεώρημα.

**Θεωρημα 5.1 (Το φίλτρο του Wiener για Ξεθολωμα).** Το βέλτιστο φίλτρο του Wiener δίνεται από, στον τομέα ανάλυσης κατά Fourier,

$$W(\omega) = \frac{K^* S_{uu}}{|K|^2 S_{uu} + S_{nn}} = \frac{K^*}{|K|^2 + r_w}, \quad (5.19)$$

όπου το ρυθμιστικό  $r_w = S_{nn}/S_{uu}$  είναι η τετράγωνισμενη, θόρυβος σε σήμα, αναλογία.

Ειδικότερα, για έναν Γκαουσιανό λευκό θόρυβο με διακύμανση  $\sigma^2$ , κάποιος έχει  $S_{nn}(\omega) \equiv \sigma^2$ . Δεδομένου ότι  $S_{uu}$  είναι συχνά οριοθετημένο, το, κατα Wiener,



ρυθμιστικό  $\tau_w$ , λοιπόν, μπορεί να οριοθετηθεί άνω του μηδενός. Το Σχήμα 5.5 δείχνει την απόδοση του φιλτραρίσματος κατά Wiener για μια θολή και μια θορυβώδη εικόνα δοκιμής.

Θα ολοκληρώσουμε αυτό το τμήμα, με δύο παρατηρήσεις σχετικά με τα κύρια χαρακτηριστικά καθώς και τα όρια του ξεθολώματος κατά Wiener.

(α) Το θόλωμα πρέπει να είναι χωρικά ομοιογενές και το PSF σαφώς γνωστό.

(β) Τόσο ο θόρυβος, όσο και η ιδανική εικόνα, πρέπει να είναι ευρείας ομοιογενούς αίσθησεως (WSH), και οι στατιστικές τους ιδιότητες να έχουν εκτιμηθεί εκ των προτέρων.

Τα ζητήματα αυτά έχουν αντιμετωπιστεί εν μέρει από ορισμένες εκδόσεις των φίλτρων του Wiener, που είναι τοπικές ή προσαρμοστικές.

#### 5.4 ΞΕΘΟΛΩΜΑ ΤΩΝ BV ΕΙΚΟΝΩΝ ΜΕ ΓΝΩΣΤΟ PSF

Σε αυτή την ενότητα, θα συζητήσουμε τον τρόπο ενίσχυσης και αποκατάστασης θολών εικόνων με BVs όταν το θόλωμα είναι γραμμικό και μετατοπιστικά αμετάβλητο, και το PSF είναι γνωστό. Η αμεταβλητή μετατόπιση μειώνει τέτοιες διαδικασίες θόλωματος στους φορείς συνέλιξης, και το ξεθολωμα ως εκ τούτου ονομάζεται επίσης, αποσυνέλιξη σε αυτό το πλαίσιο.

##### 5.4.1 Το Μεταβολικό Μοντέλο

Θεωρήστε το μοντέλο θόλωματος με μετατοπιστικά αμετάβλητο θόλωμα και πρόσθετο λευκό θόρυβο:

$$u_0(x) = k * u(x) + n(x), \quad x = (x_1, x_2) \in \mathbb{R}^2.$$

Ας υποθέσουμε ότι το PSF  $k$  είναι, είτε γνωστό σαφώς, εκτιμάται εκ των προτέρων, ή μοντελοποιήθηκε σωστά, όπως έγινε στις αρχές του τρέχοντος κεφαλαίου. Επιπλέον, στην ενότητα αυτή γίνεται η παραδοχή ότι η ιδανική εικόνα  $u$  είναι οριοθετημένη τηλεοπτικά, ή ισοδύναμα, ανήκει στο χώρο  $BV(\mathbb{R}^2)$ .

Ως πρόβλημα εκτίμησης, το ξεθολωμα θα μπορούσε να πραγματοποιηθεί υπό το φως της Μπεϋζιανής αρχής ή MAP:

$$\hat{u} = \max \text{Prob}(u | u_0, k),$$

ή ισοδύναμα, όσον αφορά την λογαριθμική ενέργεια  $E = -\log p$ ,

$$\hat{u} = \min E[u | u_0, k].$$

Ο Μπεϋζιανός τύπος με γνωστό το PSF  $k(x)$  εκφράζεται με

$$\text{Prob}(u | u_0, k) = \text{Prob}(u | k)\text{Prob}(u_0 | u, k)/\text{Prob}(u_0 | k).$$

Με δεδομένη μια εικόνας παρατήρησης  $u_0$ , ο παρονομαστής είναι απλά μια προκαθορισμένη κανονικοποίησης πιθανότητα σταθερας. Έτσι ουσιαστικά καποιος ζητά εναν εκτιμητή  $\hat{u}$ , για την ελαχιστοποίηση του προϊόντος του προηγούμενου μοντέλου  $\text{Prob}(u | k)$  και το μοντέλο δεδομένων (ή πιστότητα)  $\text{Prob}(u_0 | u, k)$ .

Δεδομένου ότι οι εικόνες και οι θαμπαδες είναι συχνά ανεξάρτητες, καποιος έχει  $\text{Prob}(u | k) = \text{Prob}(u)$ . Συνεπώς, από την άποψη της ενεργειακής διατύπωσης, καποιος προσπαθεί να ελαχιστοποιήσει την μεταγενεστερη ενέργεια

$$E[u | u_0, k] = E[u] + E[u_0 | u, k].$$

Σύμφωνα με τις υποθέσεις του ανεξάρτητου λευκού θορύβου και του προηγούμενου  $BV$ , η μεταγενεστερη ενέργεια είναι υπό τη μορφή

$$\begin{aligned} E[u | u_0, k] &= E[u] + E[u_0 | u, k] \\ &= \alpha \int_{\Omega} |Du| + \frac{\lambda}{2} \int_{\Omega} (k * u - u_0)^2 dx \end{aligned} \quad (5.20)$$

με  $x = (x_1, x_2) \in \Omega = \mathbb{R}^2$  και δύο κατάλληλα θετικά σταθμά  $\alpha$  και  $\lambda$ . Αυτό ήταν το μοντέλο αποκατάστασης που είχε προταθεί αρχικά και υπολογίστηκε απο τους Rudin,

Osher, και Fatemi, και αργότερα μελετήθηκε θεωρητικά από τους Chambolle και Lions.

Στη συνέχεια, είναι δύο σημαντικά ζητήματα αντιμετωπίζονται για πρώτη φορά - το ένα για την επιλογή των παραμέτρων στάθμισης και το άλλο για το πεπερασμένο των τομέων στις περισσότερες πρακτικές εφαρμογές.

Πρώτον, όσον αφορά την ελαχιστοποίηση της ενέργειας, μόνον η αναλογία  $r = \alpha / \lambda$  συμβάλλει στη διαδικασία λύσης. (Όμως, όπως ακριβώς και στην προβολική γεωμετρία, τοποθετώντας το  $\alpha$  και το  $\lambda$ , το ένα δίπλα στο άλλο είναι εύκολο και ευέλικτο στις εφαρμογές). Όσον αφορά την παραμετρική εκτίμηση σε στατιστικά στοιχεία, θα μπορούσε κανείς να αντιμετωπίζει το  $r$  ως άγνωστο, καθώς και να επεκτείνει την ενέργεια στο  $E [u, r | u_0, k]$  απορροφώντας κάποια προηγούμενη γνώση  $E [r]$ , σχετικά με το  $r$ .

Από την άλλη πλευρά, από την άποψη της οπτικής αντίληψης, δεν μπορεί να υπάρχει καμία καλύτερη δυνατή παράμετρος  $r$ , αφού διαφορετικοί παρατηρητές μπορούν να έχουν διαφορετικές οπτικές ευαισθησίες και ακολούθως να ζυγίζουν διαφορετικά από πρωτύτερα και με άλλο όρο πιστότητας. Επομένως, δεν θα έβλαπτε να αφήσουμε ρυθμιζόμενα τα δύο σταθμά ή την σχετική αναλογία τους  $r$ , και να επιτρέψουμε στους χρήστες να πειραματιστούν με τις δικές τους επιλογές.

Όπως εμπνευσμένο από το φίλτρο του Wiener (5.19) στην προηγούμενη ενότητα, παρ'όλα αυτά, φαίνεται φυσικό να υποθέσουμε ότι ο λόγος  $r = \alpha / \lambda$  είναι στην ίδια σειρά με τον τετραγώνισμένο λόγο σήματος προς θόρυβο,  $r_w$ . Ειδικότερα, το  $r$  θα πρέπει να είναι ανάλογο με τη διακύμανση  $\sigma^2$  του θορύβου, το οποίο είναι φυσικό, δεδομένου ότι με την Bayesian λογική, για τουλάχιστον τετραγωνικές πεποιθήσεις, έχουμε  $\lambda = O(1 / \sigma^2)$ , όπως και στις Γκαουσιανές κατανομές.

Το δεύτερο ζήτημα, αφορά το πεπερασμένο των πεδίων της εικόνας, στις περισσότερες πρακτικές εφαρμογές. Στο μοντέλο (5.20), έχει εύκολα θεωρηθεί ότι το πεδίο  $\Omega$  της εικόνας, είναι το σύνολικο επίπεδο  $\mathbb{R}^2$ , για να διευκολυνθεί η έννοια του μετατόπιστικα αμεταβλητου. Σε πραγματικές εφαρμογές, όμως, το  $\Omega$  είναι συχνά πεπερασμένο, όπως ένας δίσκος ή ένα τετράγωνο, οποτε σε αυτή την περίπτωση η θαμπάδα

$$K[u] = k * u(x) = \int_{\mathbb{R}^2} k(x-y)u(y)dy, \quad x \in \Omega,$$

πρέπει να οριστεί με σαφήνεια δεδομένου ότι το  $u$  είναι διαθέσιμο μόνο στο  $\Omega$ . Υπάρχουν δύο κοινοί τρόποι για να ξεπεραστεί αυτό το εμπόδιο. Πρώτον, πρέπει να τροποποιήσουμε την υπόθεση του αναλλοίωτου μετατόπιστικά, επιτρέποντας την μετατόπιση της διακύμανσης στο  $\Omega$ . Στη συνέχεια, ο φορέας θόλωματος  $K[u]$  παραμένει γραμμικός, αλλά μπορεί να έχει εναλλακτική μετατόπιση PSF, όπως

$$k(x, y) = \frac{k(x-y)}{\int_{\Omega} k(x-z)dz} \quad \forall x, y \in \Omega, \quad (5.21)$$

για το οποίο το γραμμικό θόλωμα αναβαθμίζεται σε

$$K[u] = \int_{\Omega} k(x, y)u(y)dy.$$

Έχει θεωρηθεί εδώ ότι το αρχικό PSF  $k(x)$  είναι μη αρνητικό, και ότι το  $x = (0,0)$  ανήκει στην ενίσχυση της πράξης  $d\mu(x) = k(x) dx$ . Δηλαδή, το ολοκλήρωμα του  $k(x)$  σε κάθε περιοχή του  $(0,0)$  είναι θετικό. Ακόμη και το μοντέλο θόλωματος κίνησης στην (5.4), για παράδειγμα, ικανοποιεί αυτόν τον όρο, του οποίου, το  $k$  είναι μια γενικευμένη συνάρτηση που περιλαμβάνει τον συντελεστή δέλτα του Dirac. Στη συνέχεια, ο παρονομαστής του ορισμού (5.21) παραμένει πάντοτε μη μηδενικός. Είναι ασήμαντο να δούμε ότι το τροποποιημένο μοντέλο θόλωματος  $K$  ή το  $k(x, y)$  όπως στο (5.21) ικανοποιεί πράγματι την υποθεση DC (5.11),  $K[1] = 1$ .

Η δεύτερη προσέγγιση για το χειρισμό ενός πεπερασμένου πεδίου  $\Omega$ , είναι η επέκταση του  $u$  πέρα από το  $\Omega$ , αντί της τροποποίησης του μετατόπιστικά αμετάβλητου πυρήνα στο  $\Omega$ , όπως ακριβώς γίνεται παραπάνω. Ας αφήσουμε την

$$Q : u|_{\Omega} \rightarrow \tilde{u} = Q[u]|_{\mathbb{R}^2}$$

να είναι ο κατάλληλος γραμμικός χειριστής προεκτασης, που επεκτείνει το  $u$  στο  $\Omega$ , σε ολόκληρο το επίπεδο. Στη συνέχεια, κανείς μπορεί να τροποποιήσει το θόλωμα σε

$$K[u](x) = k * \tilde{u}(x) = k * Q[u](x) \quad \forall x \in \Omega, \quad (5.22)$$

ή ισοδυναμια,  $K = 1_{\Omega} \cdot (k * Q)$ , όπου το  $1_{\Omega}(x)$  αντιμετωπίζεται ως πολλαπλασιαστής.

Η υπόθεση DC ικανοποιείται αν και μόνο αν  $k * Q[1] \equiv 1$  όταν περιορίζεται στο  $\Omega$ . Ειδικότερα, η φυσική προϋπόθεση  $Q[1] \equiv 1$  θα αρκούσε, δεδομένου ότι το  $k$  ικανοποιεί την DC υποθεση στο  $\mathbb{R}^2$ .

Γενικότερα, ας υποθέσουμε ότι το  $Q$  εκπροσωπείται από κάποιο πυρήνα  $g(x, y)$  με  $y \in \Omega$  και  $x \in \mathbb{R}^2$ . Τότε το τροποποιημένο  $K$  εκφράζεται με

$$k(x, y) = \int_{\mathbb{R}^2} k(x - z)g(z, y)dz, \quad x, y \in \Omega.$$

Κατά συνέπεια, η υπόθεση DC ικανοποιείται όταν το  $g$  και το  $k$  πληρούν τον όρο συμβατότητας

$$\int_{\Omega} \int_{\mathbb{R}^2} k(x - z)g(z, y)dzdy \equiv 1 \quad \forall x \in \Omega. \quad (5.23)$$

Μια άλλη λιγότερο παραδοσιακή προσέγγιση για το χειρισμό ενός πεπερασμένου τομέα, βασίζεται στην τεχνική αναδομησης της εικονας (inpainting). Έστω ότι το  $k(x)$  υποστηρίζεται σε ένα δίσκο  $B_{\rho}(0) = \{x \in \mathbb{R}^2 \mid |x| < \rho\}$  για κάποια θετική ακτίνα  $\rho$ . Ορίζουμε την  $\rho$ -περιοχή του  $\Omega$ , με

$$\Omega_{\rho} = \{x \in \mathbb{R}^2 \mid \text{dist}(x, \Omega) < \rho\}.$$

Ας υποθέσουμε ότι η ιδανική εικόνα  $u \in BV(\Omega)$  και, αντί του αρχικού μοντέλου (5.20), κανείς προσπαθεί να ελαχιστοποιήσει

$$E[u \mid u_0, k] = \alpha \int_{\Omega_{\rho}} |Du| + \frac{\lambda}{2} \int_{\Omega} (k * u - u_0)^2 dx. \quad (5.24)$$

Τώρα, η συνέλιξη μέσα στον όρο πιστότητας δεν υποδαυλίζει κανένα πρόβλημα. Θα μπορούσαμε να ονομάσουμε αυτή την προσέγγιση, οριακού στρώματος προσέγγιση, δεδομένου ότι ένα στρώμα πάχους  $\rho$  είναι κολλημένο στα όρια του  $\Omega$ .

Εν κατακλείδι, τόσο η, περιορισμένου πυρήνα, μέθοδος (5.21), όσο και η μέθοδος προεκτασης της (5.22), οδηγούν σε μετατόπιστικά εναλλακτική θαμπάδα  $K$ , με πυρήνα  $k(x, y)$ , και έτσι το μοντέλο ξεθολωματος γίνεται

$$\min_u E[u \mid u_0, K] = \alpha \int_{\Omega} |Du| + \frac{\lambda}{2} \int_{\Omega} (K[u] - u_0)^2 dx. \quad (5.25)$$

### 5.4.2 Υπαρξη και Μοναδικότητα

Υποθετούμε ότι η εικόνα του τομέα  $\Omega$  αποτελεί έναν οριοθετημένο Lipschitz τομέα, στο  $\mathbb{R}^2$ . Επιπροσθετα, έχουμε τις ακόλουθες υποθέσεις.

Υπόθεση 1. Η ιδανική εικόνα  $u \in BV(\Omega)$ .

Υπόθεση 2. Η θολή και θορυβώδη παρατήρηση  $u_0 \in L^2(\Omega)$ .

Υπόθεση 3. Το γραμμικό θολωμα  $K : L^1(\Omega) \rightarrow L^2(\Omega)$  είναι οριοθετημένο, εμβαλλομένο και ικανοποιεί την υπόθεση DC  $K[1] \equiv 1$ .

Οι συνθήκες 1 και 2 είναι προφανώς απαραίτητες για να έχει νόημα η ενεργειακή διαμόρφωση (5.25). Η εμβαλλομένη υπόθεση, είναι επίσης αναγκαία για τη μοναδικότητα του βέλτιστου ξεθολωματος.

**Θεώρημα 5.2 (Υπαρξη και μοναδικότητα του BV Ξεθολωματος).** Σύμφωνα με τις τρεις προηγούμενες συνθήκες, η βέλτιστη εκτίμηση ξεθολωματος  $u_* = \operatorname{argmin} E[u \mid u_0, K]$  στο μοντέλο (5.25) υπάρχει και είναι μοναδική.

**Απόδειξη.** Η Μοναδικότητα είναι πιο απλή. Το μετρο TV Radon, είναι ένας ομοιογενής κανόνας, και, συνεπώς, κυρτός (όχι όμως αυστηρά). Από την άλλη πλευρά,  $\|v - g\|^2$  είναι αυστηρά κυρτό στο  $v$  για κάθε δεδομένο  $g$  στο  $L^2$ . Ως εκ τούτου,  $E[u_0 \mid u, K] = \|K[u] - u_0\|^2$  είναι κυρτό δεδομένου ότι το  $K$  είναι γραμμικό, και είναι αυστηρά κυρτό αν και μόνο αν το  $K$  είναι εμβαλλομένο, το οποίο είναι όντως μέρος της υπόθεσης 3. Σε συνδυασμό, το μοντέλο είναι αυστηρά κυρτό και ο περιοριστικός παραγοντας του, πρέπει να είναι μοναδικός, αν υπάρχει πραγματικά. Για το μέρος της ύπαρξης, θεωρήστε κάθε ελαχιστοποιήσιμη αλληλουχία  $(u_n \mid n = 1, 2, \dots)$  του  $E[u \mid u_0, K]$ , που πρέπει να υφίσταται, δεδομένου ότι, για παράδειγμα,  $E[u \mid u_0, K] < \infty$ . Τότε τα TV τους

$$\text{TV}(u_n) = \int_{\Omega} |Du_n|, \quad n = 1, 2, \dots$$

είναι ομοιόμορφα οριοθετημένα. Από την ανισότητα του Poincare σε δύο διαστάσεις

$$\int_{\Omega} |u_n - \langle u_n \rangle| dx \leq C(\Omega) \int_{\Omega} |Du_n| \quad \text{for some constant } C, \quad (5.26)$$

καταλήγουμε στο συμπέρασμα ότι το  $g_n = u_n - \langle u_n \rangle$  είναι οριοθετημένο στο  $L^1(\Omega)$ .

Αφού το θολωμά  $K: L^1 \rightarrow L^2$  είναι συνεχές,

$$K[g_n] = K[u_n] - \langle u_n \rangle \quad (\text{από την υποθεση DC})$$

Πρέπει να οριοθετείται στο  $L^2$ . Αλλά, το  $K[u_n]$  είναι οριοθετημένο σύμφωνα με την Υποθεση 2 και η οριοθέτηση του  $E[u_n | u_0, K]$ , υποδεικνύει ότι το  $\langle u_n \rangle$  πρέπει να είναι μια οριοθετημένη ακολουθία.

Από την ανισότητα του Poincare (5.26) και πάλι, αυτό οδηγεί στην οριοθέτηση του  $u_n$  στο  $L^1(\Omega)$  και ως εκ τούτου στο BV  $(\Omega)$ . Με την προ-πυκνότητα  $L^1$  ενός οριοθετημένου BV-σέτ, υπάρχει μια υποακολουθία, για την ευκολία συμβολίζεται με  $(u_n | n = 1, 2, \dots)$ , η οποία συγκλίνει σε κάποιο  $u_*$  στο  $L^1(\Omega)$ . Με τη χαμηλότερη ημισυνοχή του μέτρου TV Radon υπό την  $L^1$ -τοπολογία,

$$\int_{\Omega} |Du_*| \leq \liminf_{n \rightarrow \infty} \int_{\Omega} |Du_n|. \quad (5.27)$$

Από την άλλη πλευρά, αφού το  $K: L^1(\Omega) \rightarrow L^2(\Omega)$  είναι συνεχές, πρέπει να έχουμε

$$\int_{\Omega} (K[u_*] - u_0)^2 dx = \lim_{n \rightarrow \infty} \int_{\Omega} (K[u_n] - u_0)^2 dx. \quad (5.28)$$

Με τον συνδυασμό του (5.27) και του (5.28), καταλήγουμε στο συμπέρασμα ότι

$$E[u_* | u_0, K] \leq \liminf E[u_n | u_0, K], \quad \text{και συνεπώς } u_* \text{ πρέπει}$$

αν είναι ένας περιοριστικός παραγοντας, ο οποίος συμπληρώνει την αποδείξη.

**Πόρισμα 5.3 (Ο μέσος περιορισμός).** Μετά τις υποθέσεις του προηγούμενου θεωρήματος, ο μοναδικός περιοριστικός παραγοντας  $u_*$  πρέπει να πληροί αυτομάτως τον μέσο περιορισμό  $\langle K[u_*] \rangle = \langle u_0 \rangle$ .



Στατιστικά, αυτό είναι ένα φυσικό συμπέρασμα από το μοντέλο θολώματος

$$u_0 = K[u] + n$$

δεδομένου ότι το  $n$  θόρυβος έχει μηδενική μέσο ορο.

**Αποδειξη.** Βασιζόμενοι στον μοναδικό περιοριστικό παραγοντα  $u_*$ , καθορίζουμε έναν ενιαίο μεταβλητή λειτουργίας του  $c \in \mathbb{R}$  από

$$e(c) = E[u_* - c \mid u_0, K].$$

Τότε, το  $c_* = \operatorname{argmin}_c e(c)$  πρέπει να ελαττώσει

$$\int_{\Omega} (K[u_*] - u_0 - c)^2 dx \quad \text{since } K[c] = c.$$

Ως αποτέλεσμα, ο περιοριστικός παραγοντας  $c_* = \langle K[u_*] - u_0 \rangle$ . Από την άλλη πλευρά, το  $c_*$  πρέπει να είναι μηδενικό εφόσον  $u_* - c_* = u_*$  λόγω της μοναδικότητας. Έτσι,

$$\langle K[u_*] \rangle = \langle u_0 \rangle,$$

που είναι ο ισχυρισμός.



an out-of-focus image



deblurred image (with known PSF)



Εικόνα

5.6. Αποβολή θολώματος από μια εικόνα εκτός εστίασης (με γνωστό PSF).

Η μη γραμμική εκφυλισμένη ελλειπτική εξίσωση (5.29) διακανονίζεται συχνά στο

$$\alpha \nabla \cdot \left[ \frac{\nabla u}{|\nabla u|_a} \right] - \lambda K^*[K[u] - u_0] = 0, \quad (5.30)$$

όπου η

εγγραφή  $|x|_a$  δηλώνει  $\sqrt{x^2 + a^2}$  για κάποια σταθερή θετική παράμετρο  $\alpha$ . Αυτή η διακανονισμένη εκδοχή ανταποκρίνεται στην ελαχιστοποίηση του τροποποιημένου μοντέλου αποβολής θολώματος

$$E_a[u | u_0, K] = \alpha \int_{\Omega} \sqrt{|Du|^2 + a^2} + \frac{\lambda}{2} \int_{\Omega} (K[u] - u_0)^2 dx, \quad (5.31)$$

το

οποίο είναι στενά συνδεδεμένο με το πρόβλημα ελάχιστης επιφάνειας [137]. Προσέξτε ότι η μέτρηση της περιοχής επιφάνειας στο πρώτο χρονικό κομμάτι καθορίζεται με την διανεμητική έννοια ως όμοια με τη τηλεοπτική μέτρηση ραδονίου [117, 137].

Η διακανονισμένη μη γραμμική εξίσωση (5.30) λύνεται συχνά από κάποιο επαναληπτικό σύστημα μετά από μια φυσική διαδικασία γραμμικοποίησης. Για παράδειγμα, το πιο γνωστό σύστημα,

βασίζεται στον καλύτερο παρόν υπολογισμό  $u^{(n)}$  για να επιλύσει για το  $u^{(n+1)}$  την ακόλουθη γραμμικοποιημένη εξίσωση:

$$\alpha \nabla \cdot \left[ \frac{\nabla u^{(n+1)}}{|\nabla u^{(n)}|_a} \right] - \lambda K^*[K[u^{(n+1)}] - u_0] = 0 \quad (5.32)$$

με την οριακή συνθήκη Νούμαν. Προσέξτε ότι δεδομένου του  $u^{(n)}$ , ο γραμμικός χειριστής

$$L_n = -\alpha \nabla \cdot \frac{1}{|\nabla u^{(n)}|_a} \nabla + \lambda K^* K$$

είναι ορισμένος θετικά. Αυτή είναι η φυσική έκταση του αλγόριθμου με χρονική υστέρηση (διάχυση) σε σταθερό σημείο [1,56,62] όπως συζητήθηκε προηγουμένως στο Κεφάλαιο 4.5.5.

Οι Εικόνες 5.6, 5.7 και 5.8 αποδουκνύει την εκτέλεση του πιο πάνω μοντέλου και αλγόριθμου σε τρία τυπικά παραδείγματα.

image blurred by horizontal hand jittering



deblurred image (with known PSF)



**Εικόνα**

**5.7.** Αποβολή θολώματος από μια θολωμένη εικόνα λόγω οριζόντιου τρέμουλου χεριού.

image blurred by horizontal hand jittering



deblurred image (with known PSF)



**Εικόνα**

**5.8.** Άλλο ένα παράδειγμα επαναφοράς μιας θολωμένης εικόνας λόγω κίνησης.

## 5.5 Μεταβλητή Τυφλή Αποβολή Θολώματος με Άγνωστο PSF

Σε αυτό το κεφάλαιο, αναπτύσσουμε μεταβλητά μοντέλα αποβολής θολώματος όταν το το θόλωμα PSF είναι άγνωστο, ένα σενάριο που συχνά στη λογοτεχνία επονομαζείται ως «τυφλή αποβολή θολώματος» [124, 188, 326]. Σαν ανάλογα στην θεωρία της στατιστικής εκτίμησης, θα πρέπει να ταξινομήσουμε τέτοια μοντέλα σε αυτά που είναι *παραμετρικά* ή *μη παραμετρικά* ή *ισοδύναμα μερικώς τυφλά* ή *εξολοκλήρου τυφλά*.

### 5.5.1 Παραμετρικά Τυφλή Αποβολή Θολώματος

Πρώτα ας θεωρήσουμε το σενάριο όταν το άγνωστο γραμμικό θόλωμα ανήκει σε μια γνωστή παραμετρική οικογένεια

$$\mathcal{K} = \{K_\theta \mid \theta \in I \subseteq \mathbb{R}^d\},$$

όπου το  $\theta = (\theta_1, \dots, \theta_d)$  δηλώνει ένα  $d$ -διαστατικό παραμετρικό φορέα και διαφέρει ανάλογα με το θέμα ή τον τομέα  $I$  in  $\mathbb{R}^d$ . Έτσι κάποιος δεν είναι εντελώς «τυφλός» στο τελεστήθολώματος και η αβεβαιότητα βρίσκεται στην επιλογή της σωστής παραμέτρου  $\theta$ . Για παράδειγμα, εάν το  $K$  είναι μια οικογένεια γραμμικών μετατοπιστικά-αναλλοίωτων θολωμάτων, τότε

κάθε  $K_\theta$  συνδέεται μοναδικά με ένα παραμετρικό PSF  $k(x \mid \theta)$ . Η Γκαουσιανή οικογένεια είναι ένα γνωστό παράδειγμα:

$$g(x \mid \theta) = \frac{1}{2\pi\theta} \exp\left(-\frac{x_1^2 + x_2^2}{2\theta}\right), \quad \theta \in I = (0, \infty), \quad (5.33)$$

όπου στη στατιστική το  $\theta$  ενσωματώνει τη διακύμανση  $\sigma^2$ .

Αν το θόλωμα δημιουργείται από ένα σύνθετο φυσικό σύστημα, η παράμετρος  $\theta$  καθορίζεται συχνά από τις φυσικές σταθερές εκείνου του συστήματος όπως είναι εκδηλο στις φόρμουλες (5.6) και (5.7) για το μοντέλο θολώματος εκτός εστίασης. Στην περίπτωση ατμοσφαιρικού θολώματος, για επιπλέον παράδειγμα, η παράμετρος θα μπορούσε να συνδέεται με τις εργασίες ανατάραξης της ατμόσφαιρας, όπως η θερμοκρασία, η υγρασία κλπ.

Για να αποβληθεί μια θολωμένη παρατήρηση με επιπρόσθετο Γκαουσιανό λευκό θόρυβο

$$u_0 = K_\theta[u] + n, \quad (5.34)$$

τόσο η ιδανική εικόνα  $u$  καθώς και η άγνωστη παράμετρος  $\theta$  πρέπει να υπολογιστούν. Σύμφωνα με τη σύντομη περιγραφή της Μπεύζιανής θεωρίας υπολογισμού στην αρχή του προηγούμενου κεφαλαίου, αποβολή θολώματος είναι η μεγιστοποίηση της οπίσθιας πιθανότητας

$$p(u, \theta \mid u_0), \quad \text{ή ισοδύναμα,} \quad p(u_0 \mid u, \theta)p(u, \theta).$$

Μιας και η ιδανική εικόνα και ο μηχανισμός θολώματος του συστήματος εικόνας είναι συνήθως

ανεξάρτητα, κάποιος έχει  $p(u, \theta) = p(u)p(\theta)$ . Έτσι από πλευράς τυποποίησης

λογαριθμικής ενέργειας  $E = -\log p$ , η αποβολή θολώματος γίνεται η ελαχιστοποίηση του

$$E[u, \theta | u_0] = E[u_0 | u, \theta] + E[u] + E[\theta]. \quad (5.35)$$

Οι πρώτοι δυο όριο μπορούν να αντιγραφούν ασφαλισμένα από το μη-τυφλό μοντέλο αποβολής θολώματος στην προηγούμενη ενότητα. Συνεπώς, είναι αρκετό να εσσωματώσουν κάποιο κατάλληλο μοντέλο για την παραμετρική κατανομή  $p(\theta)$  ή  $E[\theta]$ .

Τώρα θεωρήστε την κατάσταση όταν  $u \in \text{BV}(\Omega), \theta \in I \subseteq \mathbb{R}^d$ ,

για κάποια κατάλληλη λειτουργία  $\phi$ . Τότε το μοντέλο αποβολής θολώματος δίνεται αποκλειστικά από το

$$E[u, \theta | u_0] = \alpha \int_{\Omega} |Du| + \frac{\lambda}{2} \int_{\Omega} (K_{\theta}[u] - u_0)^2 dx + \phi(\theta). \quad (5.36)$$

Ας υποθέσουμε ότι το  $\phi(\theta)$  είναι περιορισμένο κάτω από το

$$\phi(\theta) \geq M > -\infty \text{ for all } \theta \in I.$$

Αλλιώς μπορεί να εξαλύψει το ρόλο των δυο πρώτων όρων στο (5.36) και να διαστρεβλώσει την πραγματική πρόθεση του μοντέλου. Σαν παράδειγμα ας θεωρήσουμε την Γκαουσιανή οικογένεια στο (5.33). Ας υποθέσουμε ότι η διακύμανση  $\theta$  υπόκειται στην εκθετική κατανομή με τη λειτουργία πυκνότητας

$$p(\theta) = a \exp(-a\theta), \quad \theta \in I = (0, \infty) \text{ for some } a > 0. \quad (5.37)$$

Τότε

$$\phi(\theta) = E[\theta] = -\ln p(\theta) = a\theta + \text{const.},$$

το οποίο είναι ξεκάθαρα περιορισμένο κάτω.

Ακολουθώντας το Θεώρημα 5.2 στην προηγούμενη ενότητα, θα έπρεπε κανονικά να υποτεθεί ότι για κάθε  $\theta \in I, K_{\theta}$  είναι εμβολικό και ικανοποιεί τη συνθήκη DC,  $K_{\theta}[1] = 1$ . Τότε για οποιοδήποτε δεδομένο  $\theta$ , ο εξαρτώμενος ελαχιστοποιητής

$$\hat{u}_{\theta} = \operatorname{argmin} E[u | u_0, K_{\theta}] = \operatorname{argmin} \alpha \int_{\Omega} |Du| + \frac{\lambda}{2} \int_{\Omega} (K_{\theta}[u] - u_0)^2 dx \quad (5.38)$$

υπάρχει

πάντα και είναι μοναδικός με βάση το Θεώρημα 5.2.

Με τον εξαρτώμενο ελαχιστοποιητή, το πρωτότυπο μοντέλο (5.36) μειώνεται σε ένα πρόβλημα

βελτιστοποίησης στον τομέα παραμέτρου

$$I \subseteq \mathbb{R}^d:$$

$$\min_{\theta \in I} e(\theta) \text{ with } e(\theta) = E[\hat{u}_{\theta}, \theta | u_0].$$

Η κυρτότητα του  $(\theta_*, \hat{u}_{\theta_*})$  μπορεί να είναι μη-μοναδικός.

Η ύπαρξη λύσεων το παραμετρικό μοντέλο αποβολής θολώματος (5.36) θα μελετηθεί στην επόμενη ενότητα σε ένα πιο γενικό σκητικό (βλέπε Θεώρημα 5.6). Μεγάλο μέρος της τυποποίησης και της μαθηματικής ανάλυσης στη συνεχή ρύθμιση είναι καινούριο στη λογοτεχνία, όσο είναι εις γνώση των συγγραφέων.

Υπολογιστικά, υπάρχει μια πιο διαδεδομένη προσέγγιση για την ελαχιστοποίηση

$$E[u, \theta | u_0],$$

η οποία αποκαλείται συχνά *εναλλασσόμενη ελαχιστοποίηση* ή η προσέγγιση *τεθλασμένου αστέρα*. Εμφανίζεται τακτικά στη λογοτεχνία περί επεξεργασίας εικόνας όταν οι στόχοι περιλαμβάνουν δυο ή περισσότερα άγνωστα, π.χ. το μοντέλο κατάτμησης Μάμφορτ-Σα που περιλαμβάνει την εικόνα  $u$  καθώς και το άγνωστο σύνολο άκρης  $\Gamma$  (λειτουργία χαρακτηριστικού άκρου  $z$  στον υπολογισμό της σύγκλισης  $\Gamma$ ) [11, 12, 76, 77, 116, 226], και το πρόβλημα αποβολής του τρέμουλου εικόνας περιλαμβάνει τόσο την εικόνα  $u$  όσο και τα άγνωστα οριζόντια τρέμουλα  $s$  [276].

Για να προχωρήσουμε κάποιος ξεκινά με μια αρχική υπόθεση  $\theta^{(0)}$ , το οποίο θα μπορούσε να παρθεί από το επιχείρημα του ελάχιστου  $\phi(\theta)$  για παράδειγμα. Τότε κάποιος επιτυχώς βρίσκει την εναλλασσόμενη σειρά (ανάμεσα σε  $u$  και  $\theta$ ) των εξαρτόμενων ελαχιστοποιητών

$$\theta^{(0)} \rightarrow u^{(0)} \rightarrow \theta^{(1)} \rightarrow u^{(1)} \rightarrow \dots \quad (5.39)$$

λύοντα

$$n = 0, 1, \dots,$$

ς, για

$$u^{(n)} = \operatorname{argmin} E[u | u_0, \theta^{(n)}], \text{ ακολουθούμενο από}$$

$$\theta^{(n+1)} = \operatorname{argmin} E[\theta | u_0, u^{(n)}], \text{ όπου}$$

$$E[\theta | u_0, u] = \frac{\lambda}{2} \int_{\Omega} (K_{\theta}[u] - u_0)^2 dx + \phi(\theta).$$

Σημειώ

στε την ακολουθία του Μάρκοβ περί ιδιότητας του τεθλασμένου αστέρα (5.39), π.χ., στη γλώσσα των εξαρτώμενων πιθανοτήτων

$$\operatorname{Prob}(\theta^{(n+1)} | u^{(n)}, \theta^{(n)}, u^{(n-1)}, \dots) = \operatorname{Prob}(\theta^{(n+1)} | u^{(n)}),$$

$$\operatorname{Prob}(u^{(n)} | \theta^{(n)}, u^{(n-1)}, \theta^{(n-1)}, \dots) = \operatorname{Prob}(u^{(n)} | \theta^{(n)}).$$

Επίσης

σημειώστε ότι το βήμα  $\theta^{(n)} \rightarrow u^{(n)}$  είναι μοναδικό ακολουθώντας το Θεώρημα 5.2, καθώς το

$$u^{(n)} \rightarrow \theta^{(n+1)}$$

βήμα εξαρτώμενης υπολογιστικής παραμέτρου μπορεί να είναι μη-μοναδικό. Κάποιος θα μπορούσε να ενδυναμώσει τη μοναδικότητα από κάποιου είδους σύστημα επιλογής, π.χ.,

$$\theta^{(n+1)} = \operatorname{argmin} \{ \phi(\theta) \mid \theta \in \operatorname{argmin} E[\theta \mid u_0, u^{(n)}] \},$$

με την

προϋπόθεση ότι το  $\phi(\theta)$  είναι αυστηρώς κυρτό.

$$n \geq 0,$$

**Θεώρημα 5.4 (Η Εναλλασσόμενη Ελαχιστοποίηση Είναι Μονότονη).** Για κάθε

**Απόδειξη.** Εναπόκειται στον αναγνώστη να επιβεβαιώσει με ευκολία ότι

$$E[u^{(n+1)}, \theta^{(n+1)} \mid u_0] \leq E[u^{(n)}, \theta^{(n)} \mid u_0].$$

Έστω ότι  $B(L^1, L^2) = B(L^1(\Omega), L^2(\Omega))$  δηλώνει το διάστημα του Μπανάχ

για όλους τους περιοσμένους γραμμικούς χειριστές από το  $L^1(\Omega)$  to  $L^2(\Omega)$  στο  $L^1(\Omega)$  το  $L^2(\Omega)$  στο  $L^1(\Omega)$  το  $L^2(\Omega)$  χρησιμοποιημένο με το πρότυπο χειριστή.

**Θεώρημα 5.5 (Σύγκλιση Εναλλασσόμενης Ελαχιστοποίησης).** Υποθέστε ότι

(α) η παραμετροποίηση του θολώματος

$$K : I \subseteq \mathbb{R}^d \rightarrow B(L^1, L^2), \quad \theta \rightarrow K_\theta$$

είναι μια συνεχόμενη απεικόνιση και

(β)  $\phi(\theta)$  είναι χαμηλότερα ημισυνεχόμενο στο  $\theta \in I$ . Τότε, αν όπως και στο  $n \rightarrow \infty$ .

$$u^{(n)} \rightarrow u_* \text{ in } L^1(\Omega) \text{ and } \theta^{(n)} \rightarrow \theta_* \in I,$$

το οριακό ζευγάρι  $(u_*, \theta_*)$  πρέπει να ικανοποιεί

$$u_* = \operatorname{argmin} E[u \mid u_0, \theta_*],$$

$$\theta_* = \operatorname{argmin} E[\theta \mid u_0, u_*].$$

**Απόδειξη.** Για οποιαδήποτε δυο αποδεκτά ζευγάρια  $(u, \theta)$  και  $(u_*, \theta_*)$  στο

$$\mathbf{BV}(\Omega) \times I,$$

$$\begin{aligned} \|K_\theta[u] - K_{\theta_*}[u_*]\|_2 &\leq \|K_\theta[u] - K_{\theta_*}[u]\|_2 + \|K_{\theta_*}[u] - K_{\theta_*}[u_*]\|_2 \\ &\leq \|K_\theta - K_{\theta_*}\| \times \|u\|_1 + \|K_{\theta_*}\| \times \|u - u_*\|_1. \end{aligned}$$

Συνεπώς,

σαν  $\theta \rightarrow \theta_*$  και  $u \rightarrow u_*$  στο  $L^1$ ,  $K_\theta[u] \rightarrow K_{\theta_*}[u_*]$  in  $L^2$ .

Από τον αλγόριθμο εναλλασσόμενης ελαχιστοποίησης (5.40), για κάθε  $n$ ,

$$E[u^{(n)} \mid u_0, \theta^{(n)}] \leq E[u \mid u_0, \theta^{(n)}] \quad \forall u \in \mathbf{BV}(\Omega).$$

Με

άλλα λόγια,

$$\alpha \int_\Omega |Du^{(n)}| + \frac{\lambda}{2} \|K_{\theta^{(n)}}[u^{(n)}] - u_0\|_2^2 \leq \alpha \int_\Omega |Du| + \frac{\lambda}{2} \|K_{\theta^{(n)}}[u] - u_0\|_2^2.$$

Σε

συνδυασμό με τη χαμηλότερη ημισυνέχεια της τηλεοπτικής μέτρησης στην τοπολογία  $L^1$  και την ιδιότητα συνέχειας που μόλις συμπεράσθηκε στην προηγούμενη παράγραφο, αυτό μας δίνει, σαν  $n \rightarrow \infty$ ,

$$\alpha \int_\Omega |Du_*| + \frac{\lambda}{2} \|K_{\theta_*}[u_*] - u_0\|_2^2 \leq \alpha \int_\Omega |Du| + \frac{\lambda}{2} \|K_{\theta_*}[u] - u_0\|_2^2,$$

ή

ισοδύναμα,  $E[u_* \mid u_0, \theta_*] \leq E[u \mid u_0, \theta_*]$  για οποιοδήποτε

$u \in \mathbf{BV}(\Omega)$ . Αυτό αποδुकνύει την πρώτη ταυτότητα στο (5.41). Από τη συνθήκη χαμηλότερης ημισυνέχειας στο  $\varphi$ , η δεύτερη ταυτότητα μπορεί να αποδειχτεί με τον ίδιο τρόπο.

Προσέξτε ότι η συνέχεια στην παραμετροποίηση θολώματος είναι δυνατή αλλά όχι απύθμενη.

Θεωρείστε την αναλλοίωτη μετατόπιση της Γκαουσιανής οικογένειας (5.33) στο  $\Omega = \mathbb{R}^2$  για παράδειγμα. Από την ανισότητα του Γιανγκ [193], κάποιος έχει

$$\|(K_\theta - K_{\theta'})[u]\|_2 = \|(g(x | \theta) - g(x | \theta')) * u\|_2 \leq \|g(x | \theta) - g(x | \theta')\|_2 \|u\|_1.$$

ως,

$$\|K_\theta - K_{\theta'}\| \leq \|g(x | \theta) - g(x | \theta')\|_2.$$

Για την Γκαουσιανή

$$\theta' > 0,$$

οικογένεια, το τελευταίο συγκλίνει προφανώς στο μηδέν για οποιοδήποτε και

$$\theta \rightarrow \theta'.$$

Από πλευράς των πρώτων επίσημων εναλλαγών,

$$\begin{aligned} \frac{\partial}{\partial u} E[u_* | u_0, \theta_*] &= \frac{\partial}{\partial u} E[u_*, \theta_* | u_0], \\ \frac{\partial}{\partial \theta} E[\theta_* | u_0, u_*] &= \frac{\partial}{\partial \theta} E[u_*, \theta_* | u_0]. \end{aligned}$$

Έτσι το όριο  $(u_*, \theta_*)$  ικανοποιεί όντως το σύστημα ισορροπίας εξισώσεων του αρχικού μοντέλου

αποβολής θολώματος  $E[u, \theta | u_0]$ , και συνεπώς προσφέρει ένα καλό υποψήφιο για άριστη

αποβολή θολώματος. Συγκεκριμένα, αν  $E[u, \theta | u_0]$  είναι αυστηρώς κυρτό σαν ένα

λειτουργικό του  $(u, \theta) \in BV(\Omega) \times I$ , τότε το  $(u_*, \theta_*)$  πρέπει να είναι ο μοναδικός καθολικός ελαχιστοποιητής.

### 5.5.2 Τυφλή Αποβολή Θολώματος Βασισμένη σε Παραμετρικό Πεδίο

Τώρα ας υποθέσουμε ότι το PSF  $k(x, y)$  ενός τελεστή  $K$  γραμμικού θολώματος δίνεται υπο τη μορφή

$$k(x, y) = g(y | x, \theta(x)) \tag{5.42}$$

για

κάποια γνωστή οικογένεια λειτουργιών  $g(y | x, \theta)$  καθορισμένων για



$$x, y \in \Omega, \text{ and } \theta : \Omega \rightarrow I \subseteq \mathbb{R}^d.$$

Συνεπώς το  $\theta$  είναι εξαρτώμενο από τα μικροστοιχεία εικόνας, αντί να είναι ένας σταθερός φορέας. Για κάθε δεδομένο παραμετρικό πεδίο  $\theta$ , θα ονομάσουμε ακόμη το θόλωμα από  $K_\theta$  ή  $k_\theta$ .

Για παράδειγμα, έστω ότι  $\theta(x) = (m(x), \sigma(x)) \in \mathbb{R}^2 \times \mathbb{R}^+$ , και ορίστε

$$g(y | x, \theta(x)) = g(y | m(x), \sigma(x)) = \frac{1}{2\pi\sigma^2(x)} \exp\left(-\frac{(y - m(x))^2}{2\sigma^2(x)}\right). \quad \text{Σε κάθε}$$

μικροστοιχείο εικόνας  $x \in \Omega = \mathbb{R}^2$ , αυτό είναι μια Γκαουσιανή επικεντρωμένη στο

$m(x)$  με διακύμανση  $\sigma^2(x)$ . Ας υποθέσουμε όπως στις περισσότερες περιπτώσεις ότι

$m(x) \equiv x$  για οποιοδήποτε  $x \in \mathbb{R}^2$ . Τότε κάποιος μπορεί απλά να καθορίσει το

$\theta(x) = \sigma(x)$  και απλά να μειώσει το πιο πάνω μοντέλο παραμετρικού πεδίου σε

$$g(y | x, \theta(x)) = \frac{1}{2\pi\theta(x)} \exp\left(-\frac{(y - x)^2}{2\theta(x)}\right).$$

Φυσικά, τέτοιες οικογένειες Γκαουσιανών PSF μπορούν να μοντελοποιήσουν επιτυχώς χωρικά διαφοροποιημένα θολώματα, τα οποία προκαλούνται συχνά από ανομοιογένεια μέσων (όπως για μεγάλης κλίμακας απόμακρο έλεγχο μέσω νερού ή αέρα) ή από ανομοιομορφία ταχύτητας γρήγορα κινούμενων αντικειμένων σε μια σκηνή.

Ακολουθώντας την ίδια γραμμή λογικής όπως και για το μοντέλο αποβολής θολώματος (5.35), κάποιος επιχειρεί να ελαχιστοποιήσει την οπίσθια ενέργεια στη μορφή

$$E[u, \theta | u_0] = E[u_0 | u, \theta] + E[u] + E[\theta],$$

Δεδομένου ότι ο μηχανισμός θολώματος είναι ανεξάρτητος από τα περιεχόμενα εικόνας. Το

καινούριο εδώ είναι ότι το  $\theta = \theta(x)$  ένα άγνωστο πεδίο αντί μια άγνωστη σταθερά. Σαν αποτέλεσμα, τα χαρακτηριστικά πεδίου μπορούν να δομηθούν στο προηγούμενο μοντέλο στο  $\theta$ . Για παράδειγμα, αν το πεδίο υποτίθεται ότι διανέμεται ομαλά, τότε κάποιος θα μπορούσε να επιβάλει

$$E[\theta] = \int_{\Omega} [\beta |\nabla \theta|^2 + \phi(\theta)] dx. \quad (5.43)$$

Αν για

παράδειγμα,  $\phi(\theta) = \kappa(1 - |\theta|^2)^2$  για κάποια μεγάλη παράμετρο  $\kappa$ , τότε το

$E[\theta]$  τυχάνει να είναι στενά συνδεδεμένο στην δοξασμένη θεωρία της υπεραγωγιμότητας και μεταβατικών φάσεων των Λαντάου και Γκίνσμπεργκ [37] και έχει μελετηθεί πολύ από τη μαθηματική κοινότητα (βλέπε π.χ., Μότικα [219], Κον και Στερνμπερκ [179] και Κον και Στραγκ [180]). Γενικά,

$\phi(\theta)$  ισούται με  $p(\theta)$  από την πιθανολογική οπτική γωνία (βλεπε π.χ., (5.37)). Οδηγούμενος από πραγματικές καταστάσεις κάποιος θα μπορούσε να θεωρήσει επίσης την πιο αναμεμειγμένη, αλλά γενική περίπτωση όταν το πεδίο  $\theta(x)$  είναι μετατοπιστικά αναλλοίωτο, στην οποία περίπτωση

$$\phi = \phi(\theta(x) | x) = -\ln p(\theta(x) | x).$$

Ας υποθέσουμε ότι η ιδανική εικόνα  $u$  ανήκει στο  $BV(\Omega)$  και το παραμετρικό πεδίο  $\theta$  διαμοιράζεται ομαλά όπως διευκρινίζεται από το (5.43). Το μοντέλο αποβολής θολώματος

$E[u, \theta | u_0]$  δίνεται συγκεκριμένα από το

$$\alpha \int_{\Omega} |Du| + \frac{\lambda}{2} \int_{\Omega} (K_{\theta}[u] - u_0)^2 dx + \int_{\Omega} (\beta |\nabla \theta|^2 + \phi(\theta)) dx. \quad (5.44)$$

Εδώ το

$\theta : \Omega \rightarrow I \subseteq \mathbb{R}^d$  είναι ένα πεδίο διάστασης  $d$  το οποίο παραμετροποιεί το θόλωμα

$K_{\theta}$ .

**Θεώρημα 5.6 (Υπαρξη Αποβολής Θολώματος από το Μοντέλο (5.44)).** Ας υποθέσουμε ότι

(Α) η ιδανική εικόνα  $u \in BV(\Omega)$  και η παρατήρηση  $u_0 \in L^2(\Omega)$ ;

(Β) το παραμετρικό πεδίο  $\theta \in H^1(\Omega, I)$ ,  $\mathbb{R}^d$ ; και  $I$  είναι ένας κλειστός τομέας του

(Γ) η πιθανότητα  $\phi : I \rightarrow \mathbb{R}$  είναι χαμηλότερη ημισυνεχόμενη, περιορισμένη κάτω από

$\inf \phi(\theta) > -\infty$  και ικανοποιεί

$$A|\theta| \leq \phi(\theta) + M \leq B(1 + |\theta|)^q \quad \forall \theta \in I \quad (5.45)$$

για

ορισμένες θετικές σταθερές  $A, B, M$  και  $q \in [1, \infty)$  και

(Δ) ο χειριστής παραμετρικού θολώματος ικανοποιεί τη συνθήκη DC  $K_\theta[1] \equiv 1$  επιπρόσθετα υποθέτουμε ότι η διαδικασία παραμετροποίησης

$$K : H^1(\Omega, I) \rightarrow B(L^1, L^2), \quad \theta \rightarrow K_\theta,$$

είναι συνεχόμενη  $L^2$ -Lipschitz  $\theta, \theta' \in H^1(\Omega, I)$ ,  
π.χ. για οποιοδήποτε

$$\|K_\theta - K_{\theta'}\| \leq L \|\theta - \theta'\|_{L^2}$$

Τότε υπάρχει ο ελαχιστοποιητής στο μοντέλο αποβολή θολώματος  $E[u(x), \theta(x) | u_0]$

Πρώτα κάνουμε δυο σχόλια αναφορικά με τα κίνητρα των συνθηκών προτού προχωρήσουμε στην απόδειξη.

(1) Η συνθήκη (Γ) απαιτεί το  $\phi$  να είναι περιορισμένο κάτω. Διαφορετικά, θα υπήρχε μια μηδαμινή ελαχιστοποιητική ακολουθία  $(u_n \equiv 0, \theta_n)$  με  $\phi(\theta_n) \rightarrow -\infty$  as  $n \rightarrow \infty$ .

(2) Η συμβατότητα μεταξύ συνθηκών (B) και (Γ) είναι εγγυημένη από την ανισότητα του Σοπλόλεβ [193] σε δυο διαστάσεις:

$$\|\theta\|_{L^q} \leq C(q, \Omega) \|\theta\|_{H^1} \quad \forall q \in [1, \infty),$$

$$\theta \in H^1(\Omega, I), \int_\Omega \phi(\theta(x)) dx < \infty.$$

η οποία εγγυάται ότι για κάθε

Στο καλύτερο της γνώσης μας, η τυποποίηση και η απόδειξη του θεωρήματος είναι καινούρια στην λογοτεχνία.

**Απόδειξη του Θεωρήματος 5.6.** Πρώτα, προσέξτε ότι για κάθε

$$s \in I, E[u \equiv 0, \theta \equiv s | u_0] < \infty. \quad \text{Συνεπώς, πρέπει να υπάρχει}$$

μια ελαχιστοποιητική ακολουθία

$$(u_n, \theta_n) \in BV(\Omega) \times H^1(\Omega, I), n = 1, 2, \dots \quad \text{τέτοια που}$$

$$E[u_n, \theta_n | u_0] \rightarrow \inf E[u, \theta | u_0] > -\infty, \quad n \rightarrow \infty.$$

Από το

χαμηλότερη συνθήκη δεσμευσής του  $\varphi$  στο (5.45),

$$\int_{\Omega} (\beta |\nabla \theta_n|^2 + A|\theta_n|) dx, \quad n = 1, 2, \dots$$

πρέπει να είναι

περιορισμένο. Με βάση τις ανισότητες του Ποϊνκαρέ [3, 117, 137, 193] αυτό προτείνει ότι το

$$(\theta_n) \quad \text{είναι μια περιορισμένη ακολουθία στο } H^1(\Omega, I). \quad \text{Τότε με βάση το θεώρημα έγκλεισης}$$

του Σοπόλεβ [3, 193] υπάρχει μια υποακολουθία του  $(\theta_n)$  για λόγους απλότητας δηλώνεται ακόμη

από το  $(\theta_n)$ , τέτοιο που

$$\theta_n(x) \rightarrow \theta_*(x) \quad \text{in } L^2(\Omega), \quad n \rightarrow \infty.$$

Εφόσον το  $I$  είναι κλειστό,  $\theta_*(x) \in I$ . Από την χαμηλότερη ιδιότητα ημισυνέχειας των

κανόνων του Σοπόλεβ, το όριο  $\theta_*$  πρέπει να ανήκει στο  $H^1(\Omega, I)$ . Με πιθανότατα άλλο ένα γύρο εκλέπτυνσης υποακολουθίας, κάποιος θα μπορούσε να υποθέσει περαιτέρω ότι

$\theta_n \rightarrow \theta_*$  σχεδόν βέβαια στο  $\Omega$ . Τότε από τις ιδιότητες της χαμηλότερης ημισυνέχειας τόσο στο  $\varphi$  όσο και στους κανόνες του Σοπόλεβ, σε συνδυασμό με το λήμμα του Φατού [126, 193] για το

$f_n = \phi(\theta_n)$  (το οποίο είναι ομοιόμορφα περιορισμένο κάτω στον τετελεσμένο τομέα  $\Omega$ ), κάποιος έχει

$$\beta \int_{\Omega} |\nabla \theta_*|^2 dx + \int_{\Omega} \phi(\theta_*(x)) dx \leq \liminf_{n \rightarrow \infty} \beta \int_{\Omega} |\nabla \theta_n|^2 dx + \int_{\Omega} \phi(\theta_n(x)) dx. \quad (5.47)$$

Από

την ανισότητα του Ποϊνκαρέ

$$\int_{\Omega} |u - \langle u \rangle| dx \leq C \int_{\Omega} |Du|, \quad \text{with } C \text{ independent of } u, \quad (5.48)$$

Το

$g_n = u_n - \langle u_n \rangle$  πρέπει να είναι περιορισμένο στο  $L^1(\Omega)$ . Από την άλλη, με βάση

τη συνθήκη του Λίπσχιτς (5.46) στο  $L^1(\Omega)$ , το  $K_{\theta_n}$  πρέπει να είναι περιορισμένο στο  $B(L^1, L^2)$  εφόσον το  $(\theta_n)$  ανήκει στο  $H^1$  όπως μόλις διευκρινίστηκε πάνω. Έτσι η επόμενη ακολουθία πρέπει να είναι συνδεδεμένη στο  $L^2(\Omega)$

$$K_{\theta_n}[g_n] = K_{\theta_n}[u_n] - \langle u_n \rangle \quad \text{από τη συνθήκη DC.}$$

Από την άλλη, από το αρχικό μοντέλο,

$$\|K_{\theta_n}[u_n]\|_2 \leq \|u_0\|_2 + \sqrt{2E_n/\lambda}, \quad E_n = E[u_n, \theta_n | u_0], \quad \text{Δηλών}$$

οντας ότι το  $(K_{\theta_n}[u_n])$  είναι περιορισμένο κάτω από το  $L^2$ . Έτσι η αναλογική ακολουθία  $(\langle u_n \rangle)$  πρέπει να είναι περιορισμένη, γεγονός το οποίο ως απάντηση μέσω την ανισότητας του Ποϊνκαρέ (5.48) υπαινίσσεται ότι το  $(u_n)$  είναι μια περιορισμένη ακολουθία στο  $BV(\Omega)$ . Από

την προ-πυκνότητα των περιορισμένων ομάδων  $BV$  στο  $L^1$ , υπάρχει μια υποακολουθία, που είναι ακόμη δηλωμένη από το  $(u_n)$  για λόγους ευκολίας, και κάποιο το  $u_* \in BV(\Omega)$ , τέτοιο που

$$u_n \rightarrow u_* \quad \text{in } L^1, \quad \text{and} \quad \int_{\Omega} |Du_*| \leq \liminf_{n \rightarrow \infty} \int_{\Omega} |Du_n|. \quad (5.49) \quad \text{Για να}$$

συνοψίσουμε, πιθανότατα με αρκετούς συνεχόμενους γύρους εκλέπτυνσης υποακολουθίας εφόσον είναι απαραίτητο, έχουμε βρει μια ελαχιστοποιητική ακολουθία

$$(u_n, \theta_n) \in BV(\Omega) \times H^1(\Omega), \quad \text{τέτοια που}$$

$$(u_n, \theta_n) \rightarrow (u_*, \theta_*) \quad \text{in } L^1 \times L^2,$$

$$(u_*, \theta_*) \in BV(\Omega) \times H^1(\Omega).$$

για κάποιο

Τότε η συνθήκη Λίπσχιτς (5.47) εγγυάται απαραίτητα ότι

$$K_{\theta_n}[u_n] \rightarrow K_{\theta_*}[u_*] \quad \text{in } L^2,$$

$$(u_*, \theta_*)$$

το οποίο σε συνδυασμό τόσο με το (5.47) και το (5.49), υποδηλώνει ότι το  $(u_*, \theta_*)$  πρέπει να είναι ένας ελαχιστοποιητής.

### 5.5.3. Μη Παραμετρική Τυφλή Αποβολή Θολώματος

Ας υποθέσουμε ότι το γραμμικό θόλωμα  $K$  είναι εντελώς άγνωστο εκτός ότι

$$1. \text{ ικανοποιεί τη συνθήκη DC } K[1] = 1, \text{ και}$$

2. ικανοποιεί συγκεκριμένους χωρικούς περιορισμούς που θα χαρακτηριστούν αργότερα.

Τότε η εργασία αποβολής θολώματος είναι πιθανώς πολύ πιο προκλητική από προηγούμενες περιπτώσεις. Αντί για υπολογισμό ορισμένων παραμέτρων ή παραμετρικών πεδίων, τώρα κάποιος πρέπει να αναδομήσει ένα χειριστή.

Στη συνέχεια θα θεωρήσουμε μόνο την περίπτωση του αναλλοίωτου της μετατόπισης όταν ο τομέας

εικόνας  $\Omega$  is  $\mathbb{R}^2$  και το γραμμικό θόλωμα καθορίζεται από το

$$\text{PSF } k(x), x \in \mathbb{R}^2$$

του. Τότε ο υπολογισμός ενός τελεστήθολώματος μειώνεται σε αυτόν της λειτουργίας.

Στο Μπεύζιανό πλαίσιο, κάποιος ελαχιστοποιεί την οπίσθια ενέργεια

$$E[u, k | u_0] = E[u] + E[u_0 | u, k] + E[k],$$

υποθέτωντας ότι ο μηχανισμός θολώματος δεν εξαρτάται από τα περιεχόμενα της εικόνας.

Στην περίπτωση των εικόνων BV και του Γκαουσιανού λευκού θορύβου, κάποιος έχει

$$E[u] = \alpha \int_{\mathbb{R}^2} |Du|, \quad E[u_0 | u, k] = \frac{\lambda}{2} \int_{\mathbb{R}^2} (k * u - u_0)^2 dx.$$

Συνεπώς το κλειδί για μια επιτυχή αποβολή θολώματος βρίσκεται στην κατάλληλη μοντελοποίηση πριν από το  $E[k]$ .

Σε περιπτώσεις όπου το θόλωμα είναι γνωστό ότι είναι ομαλό, π.χ., ένα Γκαουσιανό, κάποιος μπορεί να ενδυναμώσει φυσικά την κανονικότητα του Σοπόλεβ

$$E[k] = \beta \int_{\mathbb{R}^2} |\nabla k|^2 dx.$$

Αυτό ήταν πράγματι το μοντέλο που μελετήθηκε από τους Γιου και Καβέ [326].

## Το Μοντέλο Αποβολής Θολώματος Διπλού BV των Τσαν και Γουογκ

Από την άλλη, θολώματα εν κινήσει ή θολώματα εκτός εστίασης που προέρχονται από ιδανικούς ελεύθερους φακούς διάθλασης, όπως συζητήθηκε στην αρχή αυτού του κεφαλαίου, τα PSF είναι συνήθως συμπαγώς στηριζόμενα με αιχμηρά αποκομμένα όρια. Για μια τέτοια τάξη θολωμάτων, όπως για εικόνες με αιχμηρά άκρα, η τηλεοπτική κανονικότητα φαίνεται πιο ελκυστική,

$$E[k] = \beta \int_{\mathbb{R}^2} |Dk|.$$

Αυτή η επιλογή οδηγεί στο μοντέλο τυφλής αποβολής θολώματος που μελετήθηκε πρώτα από τους Τσαν και Γουογκ [77]:

$$E[u, k | u_0] = \alpha \int_{\mathbb{R}^2} |Du| + \beta \int_{\mathbb{R}^2} |Dk| + \frac{\lambda}{2} \int_{\mathbb{R}^2} (k * u - u_0)^2 dx, \quad (5.50)$$

το οποίο για λόγους ευκολίας θα αποκαλείται σαν το *μοντέλο τυφλής αποβολής θολώματος διπλού BV*.

Για τον άπειρο τομέα  $\Omega = \mathbb{R}^2$ , η μονάδα BV ορίζεται συμβατικά ως [137, 117]

$$\|u\|_{BV} = \|u\|_{L^1(\mathbb{R}^2)} + |Du|(\mathbb{R}^2). \quad (5.51)$$

Ενώ τα πλείστα αποτελέσματα στις λειτουργίες BV στη λογοτεχνία είναι για περιορισμένους τομείς, αξίζει να δώσουμε ιδιαίτερη προσοχή στην πολυπλοκότητα που προκύπτει από τον μη-περιορισμό.

1. Κατά πρώτον, για παράδειγμα, κάποιος θα μπορούσε εύκολα να μαγειρέψει μια ακτινικά

συμμετρική ομαλή λειτουργία  $g(x) = g(r)$  που ισούται με  $r^{-2}$  as  $r = |x| = \sqrt{x_1^2 + x_2^2} \rightarrow \infty$ .

Έτσι δεν μπορεί να ανήκει στο  $L^1(\mathbb{R}^2)$ . Από την άλλη,  $|g|^2 = r^{-4}$  και

$|\nabla g| = |\partial g / \partial r| = 2r^{-3}$  as  $r \rightarrow \infty$ . Επομένως,

$$g \in L^2(\mathbb{R}^2) \text{ and } |Dg|(\mathbb{R}^2) < \infty.$$

Αυτή η κατάσταση δεν μπορεί να συμβεί σε κανένα περιορισμένο τομέα.

2. Κατά δεύτερον, τα θεωρήματα της κανονικής πυκνότητας σε περιορισμένους τομείς μπορούν να

αποτύχουν στο  $\mathbb{R}^2$ . Ας θυμηθούμε ότι σε οποιοδήποτε περιορισμένο τομέα  $\Omega$  του Λίπσχιτς, μια περιορισμένη ακολουθία από  $BV(\Omega)$  πρέπει να είναι προσυμπαγής στο  $L^1(\Omega)$  [117, 137].

Όταν  $\Omega = \mathbb{R}^2$  αυτός ο ισχυρισμός είναι λανθασμένος. Για παράδειγμα, ας πάρουμε μια συμπαγώς στηριζόμενη ομαλή λειτουργία  $\phi(x) \neq 0$  on  $\mathbb{R}^2$ , και να ορίσουμε

$$g_n(x_1, x_2) = \phi(x_1 - n, x_2), \quad n = 1, 2, \dots$$

Τότε είναι βασικό να δούμε ότι

$$\|g_n\|_{BV} \equiv \|\phi\|_{BV}, \quad n = 1, 2, \dots$$

Ως αποτέλεσμα, το  $(g_n)$  είναι μια ακολουθία περιορισμένη στο  $BV(\mathbb{R}^2)$ . Ας υποθέσουμε ότι

μια ακολουθία  $(g_{n_k})$  συγκλίνει σε κάποιο  $g_*$  στο  $L^1(\mathbb{R}^2)$ . Τότε πιθανότατα με άλλο ένα γύρο εκλέπτυνσης ακολουθίας, κάποιος θα μπορούσε να υποθέσει ότι το  $(g_{n_k})$  συγκλίνει στο  $g_*$  σχεδόν παντού. Τώρα που το  $\phi$  είναι συμπαγώς στηριζόμενο, είναι απλό να δούμε ότι

$g_n(x) \rightarrow 0$  για οποιοδήποτε  $x \in \mathbb{R}^2$ . Έτσι  $g_* \equiv 0$ . Αλλά το  $g_{n_k} \rightarrow g_* = 0$  στο  $L^1$  είναι αδύνατο εφόσον

$$\int_{\mathbb{R}^2} |g_{n_k}(x) - 0| dx \equiv \int_{\mathbb{R}^2} |\phi(x)| dx > 0, \quad n = 1, 2, \dots$$

Τώρα

πρώτα επεκτείνουμε μια ανισότητα του Ποϊνκαρέ από τους περιορισμένους τομείς στο  $\mathbb{R}^2$ .

**Θεώρημα 5.7 (Η Ανισότητα του Ποϊνκαρέ για το  $BV(\mathbb{R}^2)$ ).** Ας υποθέσουμε ότι το  $u$  ανήκει στο

$BV(\mathbb{R}^2)$  με πεπεραμένη μονάδα  $BV$  καθορισμένο όπως στο (5.51). τότε

$u \in \dot{L}^2(\mathbb{R}^2)$ , και πιο συγκεκριμένα,

$$\|u\|_{L^2(\mathbb{R}^2)} \leq C |Du|(\mathbb{R}^2)$$

για κάποια σταθερά  $C$  ανεξάρτητη από το  $u$ .



**Απόδειξη.** Από κανονικές ανισότητες του Ποϊνκαρέ [3, 137, 193] υπάρχει μια καθορισμένη σταθερά

$C$ , τέτοια που για κάθε ανοιχτό δίσκο  $B_R$  επικεντρωμένο στην αρχή, και κάθε λειτουργία  $g \in BV(B_R)$

$$\|g - \langle g \rangle_R\|_{L^2(B_R)} \leq C|Dg|(B_R) \quad \text{with} \quad \langle g \rangle_R = \frac{1}{|B_R|} \int_{B_R} g(x) dx.$$

Είναι

αξιοσημείωτο και σοβαρό ότι το  $C$  δεν εξαρτάται από το  $R$  (λόγω του δεξιού νόμου αναλογίας και στις δυο πλευρές [117, 193]).

Για οποιοδήποτε  $u \in BV(\mathbb{R}^2)$  και οποιοδήποτε  $R > 0$ , ας καθορίσουμε

$g = u|_{B_R}$ , και να εφαρμόσουμε την πιο πάνω κανονική ανισότητα:

$$\|u - \langle u \rangle_R\|_{L^2(B_R)} \leq C|Du|(B_R) \quad \forall R > 0.$$

Πρόσεξτε ότι

$$\langle u \rangle_R = \frac{1}{|B_R|} \int_{B_R} u(x) dx \quad \Rightarrow \quad |\langle u \rangle_R| \leq \frac{\|u\|_{L^1(\mathbb{R}^2)}}{|B_R|}.$$

Συνεπώς,

$$\|u\|_{L^2(B_R)} \leq C|Du|(B_R) + |\langle u \rangle_R| \times \|1\|_{L^2(B_R)} \leq C|Du|(B_R) + \frac{\|u\|_{L^1(\mathbb{R}^2)}}{|B_R|^{1/2}}.$$

Εφόσον

τόσο το  $|Du|$  and  $u^2$  θα μπορούσαν να αντιμετωπιστούν ως μη-αρνητικά μέτρα στο

$\mathbb{R}^2$  περνώντας  $R \rightarrow \infty$  αποδुकνύει το θεώρημα.

Για το μοντέλο τυφλής αποβολής θολώματος διπλού BV

$$\min E[u, k | u_0] = \alpha \int_{\mathbb{R}^2} |Du| + \beta \int_{\mathbb{R}^2} |Dk| + \frac{\lambda}{2} \int_{\mathbb{R}^2} (k * u - u_0)^2 dx, \quad (5.52)$$

επιβάλλ

ουμε τώρα τις ακόλουθες φυσικές συνθήκες.

$$u_0 \in L^2(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2).$$

Συνθήκη Α. Την παρατήρηση

$$BV(\mathbb{R}^2).$$

Συνθήκη Β. Η ιδανική εικόνα  $u$  ανήκει στο

Συνθήκη Γ. Το θόλωμα PSF  $k$  ανήκει στο  $BV(\mathbb{R}^2)$  και ικανοποιεί τη συνθήκη DC

$$\int_{\mathbb{R}^2} k(x) dx = 1.$$

Ο περιορισμός  $L^2$  στη Συνθήκη Α πηγάζει φυσικά από το μοντέλο δεδομένων στο (5.52), καθώς ο περιορισμός  $L^\infty$  ικανοποιείται από τις περισσότερες συσκευές πραγματικής απεικόνισης και είναι βολικό για μαθηματική ανάλυση.

Από την άλλη, σύμφωνα με την ανισότητα του Ποϊνκαρέ που μόλις αναπτύχθηκε στο Θεώρημα 5.7, η

$$u \in L^2(\mathbb{R}^2).$$

Συνθήκη Β υπονίσσεται ότι

Έτσι, σύμφωνα με την ανισότητα το Γιανγκ [193],

$$\|k * u\|_{L^2(\mathbb{R}^2)} \leq \|k\|_{L^1(\mathbb{R}^2)} \|u\|_{L^2(\mathbb{R}^2)}.$$

Μαζί με την υπόθεση ότι  $u_0 \in L^2(\mathbb{R}^2)$ , αυτό κάνει τα δεδομένα που ταιριάζουν με τον όρο στο (5.52) πεπερασμένα και καλά καθορισμένα.

### Για τη Μοναδικότητα της Τυφλής Αποβολής Θολώματος

Η έρευνα στα θέματα μοναδικότητας βοηθά να ξεκαθαρίσουμε περαιτέρω τη μοντελοποίηση της τυφλής αποβολής θολώματος.

Αρχικά, δείχνουμε ότι υπάρχουν αρκετές κρυμμένες συμμετρίες στο μοντέλο αποβολής θολώματος διπλού BV, όπως δηλώνεται στα επόμενα τρία θεωρήματα. Τέτοιες συμμετρίες, όπως σε πολλούς άλλους τομείς των μαθηματικών, θα μπορούσαν να οδηγήσουν στη μη-μοναδικότητα λύσεων.

**Θεώρημα 5.8 (Αβεβαιότητα Εικόνας PSF).** Ας υποθέσουμε ότι το  $(u_*, k_*)$  είναι ένας ελαχιστοποιητής στο μοντέλο αποβολής θολώματος διπλού BV (5.52) με  $(\alpha, \beta, \lambda)$  και υπό τις συνθήκες A, B και Γ. Ας υποθέσουμε επιπλέον ότι

$$m = \int_{\mathbb{R}^2} u_*(x) dx = \beta/\alpha.$$

Τότε το  $(u_+, k_+) = (mk_*, u_*/m)$  πρέπει να είναι επίσης ελαχιστοποιητής για το ίδιο σύνολο παραμέτρων και υπό τις συνθήκες A, B και Γ.

**Απόδειξη.** Είναι σημαντικό να επιβεβαιώσουμε ότι

$$E[u_+, k_+ | u_0] = E[u_*, k_* | u_0] \quad \text{και το } k_+ \text{ ικανοποιεί τη συνθήκη DC.}$$

$$a = (a_1, a_2) \in \mathbb{R}^2,$$

Από την άλλη, για οποιοδήποτε δεδομένο  $a$  ορίζουμε τον μετατοπιστικό χειριστή

$$S_a : g(x) \rightarrow S_a[g] = g(x - a)$$

$g$ .

για οποιαδήποτε μετρίσιμη λειτουργία

Τότε ακριβώς όπως οι διαφορικοί τελεστές, κάποιος έχει

$$S_a[k * u] = S_a[k] * u = k * S_a[u].$$

$$u, k \in \text{BV}(\mathbb{R}^2),$$

Από την άλλη, για οποιοδήποτε

$$\int_{\mathbb{R}^2} |DS_a[u]| = \int_{\mathbb{R}^2} |Du| \quad \text{and} \quad \int_{\mathbb{R}^2} S_a[k] dx = \int_{\mathbb{R}^2} k dx.$$

Συνεπώς το ακόλουθο θεώρημα προκύπτει φυσικά.

**Θεώρημα 5.9 (Αβεβαιότητα Διαδικής-Μετάφρασης).** Ας υποθέσουμε ότι το  $(u_*, k_*)$  είναι ένας ελαχιστοποιητής στο μοντέλο αποβολής θολώματος διπλού BV (5.52) με  $(\alpha, \beta, \lambda)$  και υπό τις συνθήκες  $A, B$  και  $\Gamma$ . Τότε για οποιοδήποτε  $a \in \mathbb{R}^2, (S_a[u], S_a[k])$  είναι επίσης ελαχιστοποιητής για το ίδιο ακριβώς μοντέλο.

Για να κερδίσουμε περαιτέρω γνώσεις σχετικά με την φύση της μη-μοναδικότητας που προκύπτει από το μοντέλο αποβολής θολώματος διπλού BV, θεωρείστε τώρα ένα πιο εύκολο αλλά στενά συνδεδεμένο μοντέλο - το διπλό μοντέλο τυφλής αποβολής θολώματος του Σοπόλεβ

$$E_2[u, k | u_0] = \frac{\alpha}{2} \int_{\mathbb{R}^2} |\nabla u|^2 dx + \frac{\beta}{2} \int_{\mathbb{R}^2} |\nabla k|^2 dx + \frac{\lambda}{2} \int_{\mathbb{R}^2} (k * u - u_0)^2 dx, \quad (5.53)$$

όπου  $H^1(\mathbb{R}^2)$  έχει υποτεθεί ότι τόσο η εικόνα όσο και το PSF ανήκουν στο διάστημα του Σοπόλεβ και ότι το  $k \in L^1(\mathbb{R}^2)$  ικανοποιεί τη συνθήκη DC  $\langle k, 1 \rangle = 1$ .

Ας ανακαλέσουμε ότι τον ενιαίο μετασχηματισμό του Φουριέ μιας λειτουργίας  $g(x)$  στο  $\mathbb{R}^2$  καθορίζεται από το

$$G(\omega) = G(\omega_1, \omega_2) = \int_{\mathbb{R}^2} g(x) e^{-i2\pi\omega \cdot x} dx.$$

Έτσι οι θεμελιώδεις ιδιότητες δίνουν

$$\int_{\mathbb{R}^2} |G(\omega)|^2 d\omega = \int_{\mathbb{R}^2} |g(x)|^2 dx \quad \text{and} \quad \int_{\mathbb{R}^2} |\nabla g(x)|^2 dx = 4\pi^2 \int_{\mathbb{R}^2} \omega^2 |G(\omega)|^2 d\omega$$

Με  $\omega^2 = |\omega|^2 = \omega_1^2 + \omega_2^2$ .

Έστω ότι  $U(\omega), K(\omega),$  και  $U_0(\omega)$  δηλώνουν τον (ενιαίο) μετασχηματισμό Φουριέ των  $u, k,$  και  $u_0,$  αντίστοιχα. Τότε ο μετασχηματισμός Φουριέ του  $k * u$  γίνεται ένα άμεσο προϊόν  $K(\omega)U(\omega).$

Συνεπώς, στον τομέα της χωρικής συχνότητας  $\omega = (\omega_1, \omega_2)$ , η διπλή ενέργεια τυφλής

αποβολής θολώματος του Σοπόλεβ  $E_2[u, k | u_0]$  γίνεται  $E_2[U, K | U_0] =$

$$2\pi^2\alpha \int_{\mathbb{R}^2} \omega^2 |U(\omega)|^2 d\omega + 2\pi^2\beta \int_{\mathbb{R}^2} \omega^2 |K(\omega)|^2 d\omega + \frac{\lambda}{2} \int_{\mathbb{R}^2} |K(\omega)U(\omega) - U_0(\omega)|^2 d\omega. \quad (5.54)$$

Η συνθήκη DC στο PSF είναι τώρα απλά  $K(0) = 1$ . Επιπλέον, στον τομέα εικόνας  $u, k$ , και  $u_0$  είναι όλα αληθινά. Συνεπώς κάποιος απαιτεί ότι τόσο το  $U$  όσο και το  $K$  ικανοποιούν τη συζυγή συνθήκη

$$\bar{U}(\omega) = U(-\omega) \text{ and } \bar{K}(\omega) = K(-\omega), \quad \omega \in \mathbb{R}^2. \quad (5.55)$$

Τώρα είμαστε έτοιμοι να δηλώσουμε ένα θεώρημα μη-μοναδικότητας το οποίο είναι πιο γενικό από το Θεώρημα 5.9.

**Θεώρημα 5.10 (Αβεβαιότητα Δυσαιδικής Φάσης).** Έστω ότι το

$$(u_*, k_*) \in H^1(\mathbb{R}^2) \times H^1(\mathbb{R}^2)$$

είναι ένας ελαχιστοποιητής στο διπλό μοντέλο τυφλής αποβολής θολώματος (5.53) με  $(\alpha, \beta, \lambda)$ .

Έστω ότι

$$\phi(\omega) : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \omega \rightarrow \phi(\omega),$$

είναι οποιοσδήποτε πραγματικός παράγοντας ομαλής φάσης που είναι περιττός:

$$\phi(-\omega) = -\phi(\omega). \quad \text{Τότε}$$

$$(u_+, k_+) =$$

$$(U_*(\omega)e^{i\phi(\omega)}, K_*(\omega)e^{-i\phi(\omega)})$$

Αντίστροφοι Μετασχηματισμοί Φουριέ του

είναι επίσης ένας ελαχιστοποιητής για το ίδιο ακριβώς μοντέλο.

**Απόδειξη.** Είναι ξεκάθαρο να επιβεβαιώσουμε στον τομέα του Φουριέ ότι

$$E[u_+, k_+ | u_0] = E[u_*, k_* | u_0],$$

και ότι τόσο το  $u_+$  όσο και το  $k_+$  είναι πραγματικά αληθινά. Επιπλέον, το  $k_+$  ικανοποιεί τη συνθήκη DC εφόσον

$$\int_{\mathbb{R}^2} k_+ dx = K_+(0) = K_*(0)e^{-i\phi(0)} = K_*(0) = 1. \quad \square$$

Πέρνον

τας το  $\phi(\omega) = a \cdot \omega = a_1 \omega_1 + a_2 \omega_2$  για οποιαδήποτε σταθερό

$a = (a_1, a_2) \in \mathbb{R}^2$ , κάποιος ανακτά την αβεβαιότητα της δυαδικής μετάφρασης που δηλώθηκε στο Θεώρημα 5.9.

Για να γίνουν οι λύσεις μοναδικές, είναι συνεπώς επιθυμητό να επιβάλουμε περαιτέρω συνθήκες για να διαλύσουμε τις πιθανές συμμετρίες. Από την άλλη, το παράδειγμα της μη-πυκνότητας σχεδιασμένο ακριβώς πάνω από το Θεώρημα 5.7 εισηγείται επίσης το κρυφό ρίσκο της μη-ύπαρξης λύσεων υπό τη συμμετρία της μετάφρασης. Έτσι, η συζήτηση συγκλίνει σε ένα ενδιαφέρον σημείο: δηλαδή, το σπάσιμο της συμμετρίας φαίνεται εποφελές για τη μελέτη της μοναδικότητας και της ύπαρξης.

### Το Θεώρημα της Ύπαρξης

Προτού ερευνήσουμε την ιδιότητα της ύπαρξης, ας ξεκαθαρίσουμε την ανισότητα του Ποϊνκαρέ στο Θεώρημα 5.7 αφήνοντας τη συνθήκη  $L^1$

$$u \in L^2(\mathbb{R}^2),$$

**Θεώρημα 5.11 (Ανισότητα του Ποϊνκαρέ).** Για οποιοδήποτε

ισχύει η ακόλουθη ανισότητα του Ποϊνκαρέ:

$$\|u\|_{L^2(\mathbb{R}^2)} \leq C |Du|(\mathbb{R}^2)$$

για κάποια σταθερά  $C$  ανεξάρτητη από το  $u$ .

Προσέξτε ότι η συνθήκη  $L^2$  αυτόματα υπονοεί ότι το  $u \in L^1_{loc}$ , και συνεπώς η μέτρηση του τηλεοπτικού Ραδονίου είναι καλά καθορισμένο αν και μπορεί να είναι άπειρο. Από την άλλη, το

πεπερασμένο της μονάδας  $L^2$  εμφανίζεται απαραίτητη λόγω των αντιπαραδειγμάτων όπως το  $u \equiv 1$ .

**Απόδειξη.** Εφόσον τόσο η μονάδα  $L^2$  όσο και η μέτρηση του τηλεοπτικού Ραδονίου είναι συνεχόμενα υπό τη διαδικασία μαλακώματος [117, 137], αρκεί να αποδείξουμε το θεώρημα όταν  $u \in L^2(\mathbb{R}^2) \cap C^1(\mathbb{R}^2)$ .

Ας καθορίσουμε μια συμπαγώς στηριζόμενη και ακτινικά συμμετρική λειτουργία  $\phi$  από το

$$\phi(x) = \begin{cases} 1, & |x| < 1, \\ 0, & |x| \geq 2, \\ 2 - |x|, & 1 \leq |x| \leq 2. \end{cases}$$

Τότε  $0 \leq \phi(x) \leq 1$  για οποιοδήποτε  $x = (x_1, x_2) \in \mathbb{R}^2$ . Για οποιαδήποτε ακτίνα  $R > 0$ , ας ορίσουμε

$$\phi_R(x) = \phi(x/R),$$

το οποίο είναι συμπαγώς στηριζόμενο.

Τότε το  $\nabla \phi_R(x)$  στηρίζεται στο  $R \leq |x| \leq 2R$ , στο οποίο ισούται, και

$$\int_{\mathbb{R}^2} |\nabla \phi_R(x)|^2 dx = 2\pi \int_R^{2R} \frac{1}{R^2} r dr = 3\pi.$$

Για οποιοδήποτε δεδομένο  $u \in L^2(\mathbb{R}^2)$ , και  $R > 0$ , ας ορίσουμε

$u_R(x) = \phi_R(x)u(x)$ . Τότε σε οποιοδήποτε  $x \in \mathbb{R}^2$ ,  $|u_R(x)|^2$  συγκλίνει μονotonικά στο  $|u(x)|^2$  as  $R \rightarrow \infty$ .

Απο το θεώρημα της μονότονης σύγκλισης [193],

$$\|u\|_{L^2(\mathbb{R}^2)} = \lim_{R \rightarrow \infty} \|u_R\|_{L^2(\mathbb{R}^2)}.$$

Εφόσον κάθε  $u_R$  είναι συμπαγώς στηριζόμενο, το Θεώρημα 5.7 υπονοεί ότι

$$\|u\|_{L^2(\mathbb{R}^2)} \leq C \liminf_{R \rightarrow \infty} |Du_R|(\mathbb{R}^2). \quad (5.56)$$

Από την

άλλη, με βάση τον κανόνα προϊόντος,

$$\nabla(\phi_R u) = \phi_R \nabla u + u \nabla \phi_R,$$

και συνεπώς,

$$\int_{\mathbb{R}^2} |\nabla u_R| dx \leq \int_{\mathbb{R}^2} |\nabla u| dx + \int_{\mathbb{R}^2} |u| |\nabla \phi_R| dx. \quad (5.57)$$

Τώρα που το  $\phi_R$  στηρίζεται στο  $R \leq |x| \leq 2R$  ή στο δακτύλιο, κάποιος έχει

$$\int_{\mathbb{R}^2} |u| |\nabla \phi_R| dx \leq \|u\|_{L^2(B_{2R} \setminus B_R)} \|\nabla \phi_R\|_{L^2(\mathbb{R}^2)} = \sqrt{3\pi} \|u\|_{L^2(B_{2R} \setminus B_R)},$$

το

οποίο συγκλίνει στο μηδέν ως  $R \rightarrow \infty$  καθώς  $u \in L^2(\mathbb{R}^2)$ . Σε συνδυασμό με το (5.56) και το (5.57), αυτό αποδεικνύει το θεώρημα.

Παρακινούμενοι από τα Θεωρήματα 5.7 και 5.11, τώρα ορίζουμε το διάστημα  $BV_2$  από το

$$(\mathbb{R}^2) = \{u \in L^2(\mathbb{R}^2) \mid |Du|(\mathbb{R}^2) < \infty\}.$$

Έπειτα, από το Θεώρημα 5.7  $BV(\mathbb{R}^2) \subseteq BV_2(\mathbb{R}^2)$ . Το μεγαλύτερο διάστημα  $BV_2$  θα παίξει ένα φυσικό ρόλο για το μοντέλο τυφλής αποβολής θολώματος που θα αναφερθεί πιο κάτω.

Τώρα είμαστε σε μια καλά πληροφορημένη θέση για να μελετήσουμε την ύπαρξη του διπλού μοντέλου τυφλής αποβολής θολώματος

$$E[u, k \mid u_0] = \alpha \int_{\mathbb{R}^2} |Du| + \beta \int_{\mathbb{R}^2} |Dk| + \frac{\lambda}{2} \int_{\mathbb{R}^2} (k * u - u_0)^2 dx. \quad (5.58)$$

Όπως

είναι εμφανές από την πιο πάνω συζήτηση καθώς είμαστε υποκινούμενοι από πραγματικές εφαρμογές, είναι σημαντικό να σπάσουμε το συμμετρικό ρόλο της εικόνας  $u$  και του PSF  $k$  στο μοντέλο. Οι ακόλουθες συνθήκες θα υποθεθούν για τη μελέτη της ύπαρξης.



$$u_0 \in L^2(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2).$$

Συνθήκη (α). Παρατήρηση

$$\text{Συνθήκη (β). Εικόνα } u \in \text{BV}_2(\mathbb{R}^2), \text{ και } \|u\|_{L^\infty} \leq \|u_0\|_{L^\infty}. \text{ Συνθήκη (γ).}$$

$$\text{PSF } k \in \text{BV}(\mathbb{R}^2), \quad \langle k, 1 \rangle = 1.$$

μη-αρνητικό, και ικανοποιεί τη συνθήκη DC

Η υπόθεση  $L^\infty$  στην παρατήρηση είναι συχνά έγκυρη στις πλείστες πραγματικές εφαρμογές εφόσον οι περισσότερες συσκευές εικονογράφησης έχουν ένα μέγιστο σημείο κορεσμού. Ο έλεγχος

στη μονάδα  $L^\infty$  του  $u$  στη Συνθήκη (β) φαίνεται επίσης φυσική εφόσον το  $u_0$  ήδη «υπερακοντίζει» λόγω του αυθόρμητου θορύβου. Είναι σαν τη μέγιστη αρχή για τις γραμμικές ελλειπτικές εξισώσεις [117] για να εγγυηθεί σταθερότητα, ή από την άποψη πληροφορίας της θεωρίας [93], είναι αντίστοιχο με το να πούμε ότι καμία ανεξάρτητα νέα πληροφορία (σε σχέση με τα δεδομένα που έχουν μελετηθεί) δεν θα έπρεπε να δημιουργηθεί στην επεξεργασία δεδομένων. Τέλος, εκτός

από την απαραίτητη συνθήκη DC, στη συνθήκη (γ), το  $k$  απαιτείται να είναι μη-αρνητικό, οδηγώντας σε αυτό που ονομάζεται κινητός μέσος όρος στην επεξεργασία σημάτων [237]. Η συνθήκη μη αρνητικότητας ικανοποιείται σίγουρα από τα πλείστα θολά μοντέλα συμπεριλαμβανομένων των θολωμάτων εν κινήσει και των θολωμάτων εκτός εστίασης που αναφέρθηκαν στην αρχή.

Οι περιορισμοί που τοποθετήθηκαν διαφορετικά στο  $u$  και το  $k$  βοηθούν στο σπάσιμο του συμμετρικού τους ρόλου στο μοντέλο. Παρόλα αυτά, είναι εύκολο να επιβεβαιώσουμε ότι ακόμα και υπό τις Συνθήκες (β) και (γ), η αβεβαιότητα της δυαδικής μετάφρασης που αναφέρθηκε στο Θεώρημα 5.9 δεν έχει καταρριφθεί ακόμη, εφόσον και τα δυο σύνολα συνθηκών παραμένουν ακόμη μεταφραστικά αναλλοίωτα.

Για να εξαφανίσουμε τη δυαδικότητα της μετάφρασης, κάποιος πρέπει να αποτρέψει να «επιπλεύσει»

τουλάχιστο ένας από τους δυο στόχους: η ιδανική εικόνα  $u$  ή το PSF  $k$ . Οι εικόνες κατά ορισμό μπορούν να είναι ουσιαστικά οποιαδήποτε λειτουργία, καθώς τα πλείστα PSF είναι πραγματικά επικεντρωμένα γύρω από την αρχή (που αντιστοιχεί στη μετακίνηση των μέσων όρων κοντά στη εγγύτητα κάθε μικροστοιχείου της εικόνας). Έτσι, γίνεται φυσικό να έχουμε το  $k$  σταθερό αντί για το  $u$ .

Στη μελέτη των Τσαν και Γουόγκ [77], επιβάλλεται η ακόλουθη κεντροσυμμετρική συνθήκη:

$$\text{Συνθήκη (δ')}. \text{ Το PSF είναι κεντροσυμμετρικό: } k(-x) = k(x).$$

Από την οπτική γωνία του γραμμικού τελεστή, αυτό αντιστοιχεί με την απαίτηση το θόλωμα

$K[u]$  να είναι Ερμιτιανό. Αριθμητικά στοιχεία στο [77] φαίνεται να εισηγούνται ότι αυτή η συνθήκη

μπορεί να οδηγήσει σταθερά σε μοναδικές λύσεις αποβολής θολώματος, ή τουλάχιστο σε μια τοπική άποψη (δηλαδή τοπικά ελάχιστα). Αλλά δεν έχει ακόμη αναπτυχθεί γενική θεωρητική υποστήριξη.

Για ότι ακολουθεί, για να ενθαρρύνουμε το PSF να είναι υψηλά συγκεντρωμένο κοντά στην αρχή, επιβάλλουμε την τελευταία συνθήκη.

$$F(x) \in L^1(\mathbb{R}^2)$$

Συνθήκη (δ). Υπάρχει μια μη αρνητική λειτουργία  $F(x)$  και κάποια θετική ακτίνα  $R > 0$ , έτσι ώστε

$$0 \leq k(x) \leq F(x) \quad \forall x \in \mathbb{R}^2 : |x| \geq R. \quad (5.59)$$

Προσέξ

τε ότι εδώ η γνώση του  $R$  και του  $F(x)$  για  $|x| \geq R$  χειρίζεται ως γνωστό. Για παράδειγμα, αν  $F(x) \equiv 0$  η Συνθήκη (δ) ανιστοιχεί στην απαίτηση ότι το  $k$  είναι συμπαγώς

$$B_R = \{x \in \mathbb{R}^2 : |x| < R\}.$$

στηριζόμενο στο δίσκο

**Θεώρημα 5.12 (Υπαρξη Τυφλής Αποβολής Θολώματος Διπλού BV).** Υπό τις Συνθήκες (α), (β), (γ) και (δ), υπάρχουν οι ελαχιστοποιητές στο μοντέλο τυφλής αποβολής θολώματος διπλού BV (5.58).

**Απόδειξη.** Έστω ότι  $R > 0$  εισάγεται όπως στη Συνθήκη (δ), και ας ορίσουμε

$$k_R(x) = \frac{1}{|B_R|} 1_{|x| < R}(x) = \frac{1}{\pi R^2} 1_{|x| < R}(x).$$

,  $k_R \geq 0$   
Τότε,  $k_R$  ικανοποιεί τη συνθήκη DC και

$$\int_{\mathbb{R}^2} |Dk_R| = \frac{2\pi R}{\pi R^2} = \frac{2}{R} < \infty.$$

Συγκεκριμμένα, για το ξεχωριστό ζευγάρι εικόνας-PSF  $(u \equiv 0, k_R)$ ,

$$E[u \equiv 0, k_R | u_0] = \beta \int_{\mathbb{R}^2} |Dk_R| + \frac{\lambda}{2} \|u_0\|_{L^2}^2 < \infty.$$

Επομέν

ως, πρέπει να υπάρξει μια ελαχιστοποιητική ακολουθία  $(u_n, k_n)$  στην ενέργεια τυφλής αποβολής θολώματος διπλού BV, που υπόκεινται στις Συνθήκες από (α) μέχρι (δ).

Με βάση την ανισότητα του Ποϊνκαρέ στο Θεώρημα 5.11, το  $(u_n)$  πρέπει να είναι μια περιορισμένη

$$L^2(\mathbb{R}^2).$$

ακολουθία στο  $\Omega$  of  $\mathbb{R}^2$ , από την ανισότητα του Σβαρτζ, Τότε σε οποιοδήποτε σταθερό περιορισμένο (Λόσχιτς) τομέα

$$\|u_n\|_{L^1(\Omega)} \leq \|u_n\|_{L^2(\Omega)} \times \sqrt{|\Omega|}.$$

Συνεπώς, απομονωμένο στο  $\Omega$ ,  $(u_n)$  είναι περιορισμένο στο  $BV(\Omega)$  για οποιοδήποτε

$\Omega$ . Τότε από την προσυμπαγότητα του  $L^1$  των περιορισμένων συνόλων των λειτουργιών BV σε περιορισμένους τομείς [137, 193] σε συνδυασμό με τη διαδικασία διαγωνικής επιλογής του Κάντορ

[193], κάποιος θα μπορούσε να βρει μια υποακολουθία του  $(u_n)$ , που για ευκολία είναι ακόμη σεσημασμένο από το  $(u_n)$ , και ορισμένα  $u_* \in L^1_{loc}(\mathbb{R}^2)$ , τέτοια που σε κάθε πεπερασμένο δίσκο

$$B_\rho = \{x \in \mathbb{R}^2 : |x| < \rho\},$$

$$u_n \rightarrow u_* \quad \text{in} \quad L^1(B_\rho).$$

Συγκεκριμένα, με πιθανότητα άλλο ένα γύρο υποακολουθίας εκλέπτυνσης, κάποιος θα μπορούσε να υποθέσει ότι

$$u_n(x) \rightarrow u_*(x), \quad \text{σχεδόν παντού στο } \mathbb{R}^2. \quad (5.60)$$

Τότε από τη χαμηλότερη ημισυνέχεια της μέτρησης του τηλεοπτικού Ραδονίου υπό την τοπολογία

$L^1_{loc}$  [137], κάποιος έχει

$$\int_{\mathbb{R}^2} |Du_*| \leq \liminf_{n \rightarrow \infty} \int_{\mathbb{R}^2} |Du_n|. \quad (5.61)$$

Εφόσον όλα αυτά τα αποτελέσματα είναι βασισμένα μόνο στον περιορισμό της μέτρησης του τηλεοπτικού

Ραδονίου, ισχύουν επίσης για την ακολουθία PSF  $(k_n)$  η οποία πρέπει να υποτεθεί από τούδε και

στο εξής. Συγκεκριμένα, υπάρχει το  $k_* \in L^1_{loc}(\mathbb{R}^2)$ , τέτοιο που

$$k_n \rightarrow k_* \quad \text{σε οποιοδήποτε } L^1(B_\rho), \quad \text{και σχεδόν σίγουρα στο } \mathbb{R}^2. \quad (5.62)$$

Επιπρόσθετα,

$$\int_{\mathbb{R}^2} |Dk_*| \leq \liminf_{n \rightarrow \infty} \int_{\mathbb{R}^2} |Dk_n|. \quad (5.63)$$

Έστω

$$M = \|u_0\|_{L^\infty(\mathbb{R}^2)}.$$

ότι

Τότε η Συνθήκη (β) υπονοεί ότι

$$|u_n(y)| \leq M, \quad n = 1, 2, \dots, \quad \forall y \in \mathbb{R}^2. \quad (5.64)$$

Για

$$x \in \mathbb{R}^2, \quad k^x(y) = k(x - y), \quad k_n^x$$

οποιαδήποτε σταθερή  $x$  ας ορίσουμε  $k^x$  και  $k_n^x$

ομοίως. Ας ορίσουμε  $k_*^x$  Τότε

$$r = R + |x|.$$

όπου το  $B_r^c = \mathbb{R}^2 \setminus B_r$  υποδηλώνει το συμπλήρωμα.

Περιορισμένη στο  $k_n^x(y) \rightarrow k_*^x(y)$  η ιδιότητα στο  $L^1(B_r)$ . (5.62) υπονοεί ότι

$$\langle k_*^x, u_n \rangle_{B_r} \rightarrow \langle k_*^x, u_* \rangle_{B_r}, \quad n \rightarrow \infty.$$

Από το θεώρημα του Λεπέσκ στο οποίο κυριαρχεί η σύγκλιση, σε συνδυασμό με το (5.60) και το (5.64),

$$\langle k_*^x, u_n \rangle_{B_r} \rightarrow \langle k_*^x, u_* \rangle_{B_r}, \quad n \rightarrow \infty.$$

Από την άλλη,

$$|\langle k_n^x, u_n \rangle_{B_r} - \langle k_*^x, u_n \rangle_{B_r}| \leq M \|k_n^x - k_*^x\|_{L^1(B_r)} \rightarrow 0, \quad n \rightarrow \infty.$$

Επομέν

ως έχουμε εδραιώσει ότι

$$y \in B_r^c \text{ (with } r = R + |x|),$$

$$|y - x| \geq |y| - |x| > r - |x| = R.$$

Στο συμπλήρωμα

Τότε με βάση τη Συνθήκη (δ),

$$|k_n^x(y)| = |k_n(x - y)| \leq F(x - y) = F^x(y)$$

$$\forall n \text{ and } y \in B_r^c.$$

Εφόσον  $F \in L^1(\mathbb{R}^2)$ , για κάθε δεδομένο  $x$ , κάποιος πρέπει να έχει

$F^x \in L^1(B_r^c)$ . Χρησιμοποιώντας το  $M F_x(y)$  ως το αυξητικό και εφαρμόζοντας το θεώρημα του Λεπέσκ στο οποίο κυριαρχεί η σύγκλιση στο  $(k_n^x(y)u_n(y))$ , πάνω στο  $B_r^c$ ,

Σε συνδυασμό με το (5.66) και το (5.67), η ταυτότητα αποσύνθεσης (5.65) προτείνει ότι

$$k_n * u_n(x) \rightarrow k_* * u_*(x) \quad \forall x \in \mathbb{R}^2.$$

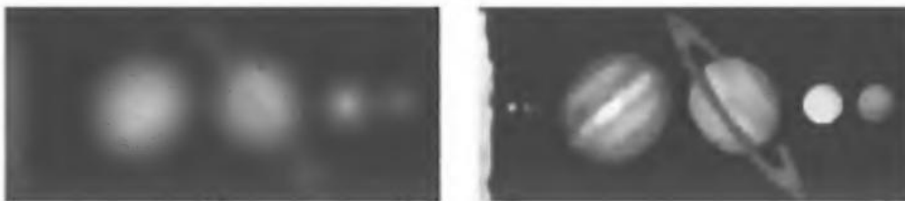
Τότε εφαρμόζοντας το λήμμα του Φατού [193] στη σημειακή συγκλινόμενη μη-αρνητική ακολουθία

$$e_n(x) = (k_n * u_n - u_0(x))^2, \quad n = 1, 2, \dots, \quad \text{κάποιος έχει}$$

$$\int_{\mathbb{R}^2} e_*(x) dx \leq \liminf_{n \rightarrow \infty} \int_{\mathbb{R}^2} e_n(x) dx.$$

Σε συνδυασμό με το (5.61) και (5.63), αυτό οδηγεί στο

$$E[u_*, k_* | u_0] \leq \liminf_{n \rightarrow \infty} E[u_n, k_n | u_0],$$



**Εικόνα**

**5.9.** Ένα υπολογιστικό παράδειγμα για το μοντέλο τυφλής αποβολής θολώματος διπλού BV (από τους Τσαν και Γουογκ [77]). Αριστερά: η θολωμένη εικόνα. Δεξιά: η αποβολή θολώματος από την εικόνα.

που δηλώνει ότι το  $(u_*, k_*)$  πρέπει να είναι ένα ελαχιστοποιητικό ζευγάρι, αν πραγματικά ικανοποιεί τις Συνθήκες (β) μέχρι (δ) οι οποίες επαληθεύονται μια προς μια ως ακολούθως.

Λόγω της σημειακής σύγκλισης, κάποιος έχει  $\|u_*\|_{L^\infty} \leq \|u_0\|_{L^\infty}$ . Επιπλέον, με βάση το λήμμα του Φατού και την ανισότητα του Ποϊνκαρέ στο Θεώρημα 5.11,

$$\|u_*\|_{L^2} \leq \liminf_{n \rightarrow \infty} \|u_n\|_{L^2} \leq \liminf_{n \rightarrow \infty} |Du_n|(\mathbb{R}^2) < \infty.$$

Αυτό

δηλώνει ότι, σε συνδυασμό με το (5.61),  $(k_n)$  και ικανοποιεί τη Συνθήκη (β).

Παρομοίως, το  $k_*$  πρέπει να ανήκει στο  $BV(\mathbb{R}^2)$ ,  $k_* \geq 0$ , και ικανοποιεί τη Συνθήκη (δ).  
 Η συνθήκη DC

$$\langle k_*, 1 \rangle = 1$$

πρέπει να ισχύει λόγω της Συνθήκης (δ) για την ελαχιστοποιητική ακολουθία  $(k_n)$  και τη σύγκλιση

$L^1$ -τους στο  $k_*$  σε οποιοδήποτε περιορισμένο τομέα του  $\mathbb{R}^2$ .

Συνεπώς, το  $(u_*, k_*)$  είναι όντως ένας ελαχιστοποιητής στο μοντέλο τυφλής αποβολής θολώματος διπλού BV (5.58), υποκείμενο στις Συνθήκες (β) μέχρι (δ).

Υπολογιστικά, το μοντέλο τυφλής αποβολής θολώματος διπλού BV (5.58) μπορεί να εφαρμοστεί μέσω του αλγόριθμου εναλλακτικής ελαχιστοποίησης (AM) (βλέπε (5.40)). Περισσότερες πληροφορίες μπορούν να βρεθούν στο [77, 229]. Η Εικόνα 5.9 είναι ένα υπολογιστικό παράδειγμα από τους Τσαν και Γουογκ [77].