

Περιεχόμενα

1	Εισαγωγή	3
1.1	Διατύπωση του προβλήματος	3
1.1.1	Κατάτμηση βάση ατλάντων	3
1.1.2	Αντικείμενο διπλωματικής εργασίας	3
2	Θεωρητικό υπόβαθρο	5
2.1	Πλαίσιο συγχώνευσης κατηγοριών βασισμένο σε γράφους	5
2.2	Ελάχιστα απόλυτος τελεστής συρρίκνωσης και επιλογή .	5
2.3	Καταχώρηση εικόνας	6
2.3.1	Μέση διαφορά τετραγώνων	7
2.3.2	Αλγόριθμος προσαρμοστικής στοχαστικής απότο- μης καθόδου	8
2.3.3	Αγχίγραμμος μετασχηματισμός	9
2.3.4	Γραμμική παρεμβολή	11
2.3.5	Χώρος κλίμακας Gauss	11
2.3.6	Δειγματοληψία συντεταγμένων εικόνας	12
2.3.7	Μάσκα εικόνας	13
2.4	Μέτρα αξιολόγησης	13
2.4.1	Συντελεστής ομοιότητας Dice	14
2.4.2	Δείκτης δομικής ομοιότητας	14

Κατάλογος σχημάτων

1	Μετατόπιση, περιστροφή, αλλαγή κλίμακας και στρέ- βλωση ενός τετραγώνου.	10
2	Χώρος κλίμακας Gauss για διάφορες τιμές του t	12

1 Εισαγωγή

1.1 Διατύπωση του προβλήματος

Η κατάτμηση είναι η διαδικασία διαμέρισης μίας εικόνας σε διάφορα ουσιαστικά τμήματα. Σκοπός της κατάτμησης είναι η απλοποίηση ή/και η αλλαγή της αναπαράστασης της εικόνας σε κάτι που είναι πιο σημασιολογικά σημαντικό και είναι πιο εύκολο να αναλυθεί. Πιο συγκεκριμένα, η κατάτμηση εικόνας είναι η διαδικασία ανάθεσης μίας ετικέτας σε κάθε εικονοστοιχείο της εικόνας έτσι ώστε τα εικονοστοιχεία με την ίδια ετικέτα να έχουν ίδια χαρακτηριστικά.

Στην ιατρική απεικόνιση αυτά τα τμήματα αντιστοιχούν συχνά σε διαφορετικές κατηγορίες ιστών, οργάνων, παθολογίες ή άλλες βιολογικά σχετιζόμενες δομές. Η αυτοματοποιημένη κατάτμηση ιατρικών απεικονίσεων μπορεί να βοηθήσει τους γιατρούς επιταχύνοντας τη διαδικασία διάγνωσης. Η καθημερινή δημιουργία πληθώρας ιατρικών απεικονίσεων καθιστά τη χειροκίνητη κατάτμηση από ειδικούς όλο και δυσκολότερη, λόγω του χρονικού διαστήματος που χρειάζεται η ανάλυση. Επομένως, η ανάπτυξη αξιόπιστων, σταθερών και ακριβών τεχνικών για την αυτοματοποιημένη κατάτμηση ιατρικών απεικονίσεων αποτελεί μία σημαντική πρόκληση.

1.1.1 Κατάτμηση βάση ατλάντων

Η κατάτμηση απεικονίσεων βάση ατλάντων αποτελεί την διαδικασία κατά την οποία χρησιμοποιούνται απεικονίσεις που έχουν κατανεμηθεί από κάποιον ειδικό, ούτως ώστε να επιτευχθεί η κατάτμηση της νέας απεικόνισης. Οι μέθοδοι αυτοί συνήθως απαιτούν την χρήση καταχώρησης εικόνας (image registration) με σκοπό την ευθυγράμμιση τού ή των ατλάντων στην εικόνα-απεικόνιση που πρόκειται να κατανεμηθεί.

1.1.2 Αντικείμενο διπλωματικής εργασίας

Στην διπλωματική εργασία εφαρμόστηκαν μέθοδοι κατάτμησης ιατρικών απεικονίσεων βάση ατλάντων με χρήση μηχανικής μάθησης. Οι απεικονίσεις αφορούν μαγνητικές τομογραφίες σε γόνατα και οι

περιοχές κατάτμησης αποτελούνται από τους αρθρικούς χόνδρους και τα οστά. Τα δεδομένα που χρησιμοποιήθηκαν προέρχονται από το Osteoarthritis Initiative Zuse Institute Berlin(OAI ZIB). Οι μέθοδοι που χρησιμοποιήθηκαν προέρχονται από μεθόδους που έχουν χρησιμοποιηθεί για την κατάτμηση περιοχών του εγκεφάλου. TODO: references

2 Θεωρητικό υπόβαθρο

2.1 Πλαίσιο συγχώνευσης κατηγοριών βασισμένο σε γράφους

Το πλαίσιο συγχώνευσης κατηγοριών βασισμένο σε γράφους προτάθηκε από το [1] ως μία μέθοδος μέσω της οποίας μπορούν να παραχθούν πολλές υπάρχουσες μέθοδοι συγχώνευσης κατηγοριών.

Έστω το ζεύγος $\{(I_i, L_i), i = 1, \dots, n\}$ όπου I_i είναι η απεικόνιση ενός άτλαντα, L_i ο χάρτης των κατηγοριών της αντίστοιχης απεικόνισης και n το σύνολο των ατλάντων. Δοθείσας μία απεικόνισης I , δημιουργείται ο γράφος G_i μεταξύ των εικονοστοιχείων \mathbf{x} της δοθείσας εικόνας I και του εικονοστοιχείου \mathbf{y} της απεικόνισης του άτλαντα I_i , μαζί με τα αντίστοιχα βάρη $w_i(\mathbf{x}, \mathbf{y})$, για $(\mathbf{x}, \mathbf{y}) \in \Omega^2$, όπου Ω ο χώρος των απεικονίσεων. Με τη δημιουργία των βαρών του γράφου, η συγχώνευση των κατηγοριών γίνεται για κάθε \mathbf{x} της απεικόνισης εισόδου I σύμφωνα με το τύπο:

$$L(\mathbf{x}) = \frac{\sum_{i=1}^n \sum_{\mathbf{y} \in \Omega} w_i(\mathbf{x}, \mathbf{y}) L_i(\mathbf{y})}{\sum_{i=1}^n \sum_{\mathbf{y} \in \Omega} w_i(\mathbf{x}, \mathbf{y})}, \forall \mathbf{x} \in \Omega \quad (1)$$

L είναι ο χάρτης των κατηγοριών της απεικόνισης εισόδου I .

2.2 Ελάχιστα απόλυτος τελεστής συρρίκνωσης και επιλογή

Ο ελάχιστα απόλυτος τελεστής συρρίκνωσης και επιλογής (least absolute shrinkage and selection operator, Lasso) [2] είναι μία μέθοδος παλινδρόμησης. Παλινδρόμηση είναι η μέθοδος πρόβλεψης της συμπεριφοράς μίας μεταβλητής βασισμένη σε μία ή περισσότερες άλλες μεταβλητές. Έστω οι παρατηρήσεις $(\mathbf{x}_i, y_i), i = 1, \dots, N$, όπου $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})$ οι p μεταβλητές εισόδου και y_i η μεταβλητή εισόδου για την i -οστή παρατήρηση. Αν $\hat{\beta} = (\hat{\beta}_1, \dots, \hat{\beta}_p)^T$ και $\hat{\alpha}$ τα αποτελέσματα της παλινδρόμησης (συντελεστές παλινδρόμησης) τότε ο τελεστής ορίζεται ως:

$$(\hat{\alpha}, \hat{\beta}) = \operatorname{argmin} \left\{ \sum_{i=1}^N \left(y_i - \alpha - \sum_{j=1}^p \beta_j x_{ij} \right)^2 \right\} \text{ subject to } \sum_{j=1}^p |\beta_j| \leq t \quad (2)$$

Το $t \geq 0$ είναι παράμετρος συντονισμού. Για κάθε τιμή του t η λύση για το α είναι $\hat{\alpha} = \bar{y}$ (όπου \bar{y} είναι η μέση τιμή του y). Επίσης μπορούμε να υποθέσουμε χωρίς την απώλεια της γενίκευσης ότι $\hat{\alpha} = 0$ ώστε να εξαλειφθεί το α . Έστω $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N)^T$ και $y = (y_1, \dots, y_N)^T$ τότε η (2) μπορεί να γραφτεί ως:

$$\hat{\beta} = \operatorname{argmin} \left\{ \frac{1}{N} \|(y - \mathbf{X}\beta)\|_2^2 \right\} \text{ subject to } \|\beta\|_1 \leq t \quad (3)$$

Όπου η p -οστή νόρμα του $x = (x_1, \dots, x_n)$ ορίζεται ως:

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}$$

Η εξίσωση (3) μπορεί να γραφτεί επίσης με την χρήση του πολλαπλασιαστή Lagrange ως:

$$\hat{\beta} = \operatorname{argmin} \left\{ \frac{1}{N} \|(y - \mathbf{X}\beta)\|_2^2 + \lambda \|\beta\|_1 \right\} \quad (4)$$

Πολλές φορές οι στήλες του \mathbf{X} κανονικοποιούνται, δηλαδή ισχύει:

$$\sum_{j=1}^p \frac{x_{ij}}{N} = 0, \quad \sum_{j=1}^p \frac{x_{ij}^2}{N} = 1$$

2.3 Καταχώρηση εικόνας

Καταχώρηση εικόνας (image registration) είναι η διαδικασία μετασχηματισμού διαφορετικών δεδομένων σε ένα σύστημα συντεταγμένων. Η διαδικασία αυτή επιδιώκει μέσω του μετασχηματισμού αυτού, την επικάλυψη των κοινών χαρακτηριστικών των δεδομένων. Τα δεδομένα

μπορεί να είναι πολλαπλές φωτογραφίες, δεδομένα από διαφορετικούς αισθητήρες, ώρες, βάθη και οπτικές [3]. Στις ιατρικές απεικονίσεις επιδιώκεται μέσω του μετασχηματισμού, αντίστοιχα εικονοστοιχεία των απεικονίσεων να αναπαριστούν όμοια βιολογικά σημεία.

Συνήθως στις μεθόδους καταχώρησης εικόνας η μία εικόνα παραμένει σταθερή κατά την διαδικασία της καταχώρησης και η άλλη-άλλες μετασχηματίζονται. Η εικόνα που παραμένει σταθερή αποκαλείται εικόνα αναφοράς ή απλά σταθερή εικόνα. Οι εικόνες που μετασχηματίζονται ονομάζονται εικόνες καταχώρισης ή κινούμενες εικόνες.

Η βασική διαδικασία της καταχώρησης εικόνας αποτελείται από την ανίχνευση χαρακτηριστικών, την ευθυγράμμιση χαρακτηριστικών από την κινούμενη εικόνα στην σταθερή, εκτίμηση παραμέτρων συναρτήσεων χαρτογράφησης, την δειγματοληψία και τον μετασχηματισμό εικόνας. Οι μέθοδοι καταχώρισης εικόνας είναι μοναδικές για κάθε πρόβλημα και δεν υπάρχει κοινή τεχνική καταχώρισης εικόνας που να είναι ισάξια αποτελεσματική σε κάθε πρόβλημα που εφαρμόζεται [4]. Για παράδειγμα υπάρχουν μέθοδοι που βασίζονται στην ένταση των εικόνων και μέθοδοι που βασίζονται σε χαρακτηριστικά τους.

2.3.1 Μέση διαφορά τετραγώνων

Στην οικογένεια των μεθόδων που βασίζονται στην ένταση των εικόνων, η μέτρηση της ομοιότητας μεταξύ της κινούμενης και της σταθερής εικόνας αποτελεί ένα δομικό στοιχείο της καταχώρησης εικόνας. Η μέτρηση αυτή αξιολογεί την καταλληλότητα του μετασχηματισμού και χρησιμοποιείται για να υπολογιστεί η παράγωγος της αξιολόγησης, ούτως ώστε να χρησιμοποιηθούν από τον αλγόριθμο βελτιστοποίησης. Η επιλογή της μετρικής βασίζεται στο εκάστοτε πρόβλημα που επιχειρεί να λύσει

Η μέση διαφορά τετραγώνων υπολογίζει τον μέσο όρο των τετραγώνων της διαφορά της έντασης για κάθε εικονοστοιχείο μεταξύ της σταθερής και της κινούμενης εικόνας. Η μετρική αυτή βασίζεται στην υπόθεση ότι η ένταση όμοιων σημείων είναι η ίδια και στις δύο εικόνες και έχει την ίδια κατανομή και στις δύο εικόνες. Αυτό σημαίνει ότι και οι δύο εικόνες έχουν την ίδια μορφή.

Η μέση διαφορά τετραγώνων ορίζεται ως:

$$MSSD(I_F, I_M) = \frac{1}{|\Omega_F|} \sum_{\mathbf{x}_i \in \Omega_F} (I_F(\mathbf{x}_i) - I_M(\mathbf{T}(\mathbf{x}_i)))^2$$

Όπου I_F η σταθερή εικόνα, I_M η κινούμενη εικόνα, \mathbf{T} ο μετασχηματισμός της διαδικασίας της καταχώρησης, $\Omega_F \subset \mathbb{R}^D \mapsto \mathbb{R}$ το πεδίο ορισμού της σταθερής εικόνα και $|\Omega_F|$ ο αριθμός των εικονοστοιχείων της σταθερής εικόνας.

2.3.2 Αλγόριθμος προσαρμοστικής στοχαστικής απότομης καθόδου

Ο Αλγόριθμος προσαρμοστικής στοχαστικής απότομης καθόδου (adaptive stochastic gradient descent, ASGD) [5] είναι ένας αλγόριθμος βελτιστοποίησης. Ο αλγόριθμος αυτός χρησιμοποιείται ώστε να υπολογιστούν οι βέλτιστες παράμετροι του μετασχηματισμού. Χρησιμοποιεί μία επαναληπτική διαδικασία ώστε να υπολογίσει τις βέλτιστες παραμέτρους.

Έστω $F(x) : \Omega_F \subset \mathbb{R}^D \mapsto \mathbb{R}$ η σταθερή εικόνα, $M(x) : \Omega_M \subset \mathbb{R}^D \mapsto \mathbb{R}$ η κινούμενη εικόνα και $\mathbf{T}(\mathbf{x}, \boldsymbol{\mu}) : \Omega_F \times \mathbb{R}^P \mapsto \Omega_M$ παραμετροποιημένος μετασχηματισμός συντεταγμένων, όπου $\boldsymbol{\mu} \in \mathbb{R}^P$ αναπαριστά το διάνυσμα των παραμέτρων του μετασχηματισμού. Το πρόβλημα βελτιστοποίησης είναι:

$$\hat{\boldsymbol{\mu}} = \underset{\boldsymbol{\mu}}{\operatorname{argmin}} \mathcal{C}(F, M \circ \mathbf{T})$$

Όπου \mathcal{C} είναι μία συνάρτηση αξιολόγησης (όπως η μέση διαφορά τετραγώνων 2.3.1). Η $\hat{\boldsymbol{\mu}}$ είναι η παράμετρος που ελαχιστοποιεί την συνάρτηση κόστους, επομένως μπορούμε να γράψουμε την συνάρτηση αυτή $\mathcal{C}(\boldsymbol{\mu}) \equiv \mathcal{C}(F, M \circ \mathbf{T})$.

Ο Αλγόριθμος προσαρμοστικής στοχαστικής απότομης καθόδου οπότε ορίζεται ως:

$$\begin{aligned} \boldsymbol{\mu}_{k+1} &= \boldsymbol{\mu}_k - \gamma(t_k) \tilde{\mathbf{g}}_k, \quad k = 0, 1, \dots, K \\ t_{k+1} &= [t_k + f(-\tilde{\mathbf{g}}_k^T \tilde{\mathbf{g}}_{k-1})]^+ \end{aligned}$$

Όπου $[x]^+$ σημαίνει $\max(x, 0)$, f είναι μία σιγμοειδής συνάρτηση και μ_0 , t_0 και t_1 είναι ορίσματα του αλγορίθμου. \tilde{g}_k είναι προσέγγιση της παραγώγου $g \equiv \frac{\partial C}{\partial \mu}$ στο μ_k . Επίσης ο βαθμωτός συντελεστής κέρδους γ_k ορίζεται από την συνάρτηση απόσβεσης:

$$\gamma_k \equiv \gamma(k) = \frac{a}{(k + A)^a}$$

Όπου $0 < a \leq 1$ και A σταθερά ορίσματα του αλγορίθμου.

Στον αλγόριθμο αυτό, η μεταβλητή του "χρόνου" t_k εξαρτάται από το εσωτερικό γινόμενο της παραγώγου \tilde{g}_k και της προηγούμενης παραγώγου \tilde{g}_{k-1} . Αν οι παράγωγοι αυτοί σε δύο συνεχόμενα βήματα του αλγορίθμου έχουν την ίδια κατεύθυνση, τότε το εσωτερικό τους γινόμενο θα είναι θετικό. Οπότε ο "χρόνος" θα μειωθεί, που οδηγεί σε μεγάλο βήμα $\gamma(t_{k+1})$ επειδή η γ είναι γνησίως φθίνουσα συνάρτηση. Με αυτό τον τρόπο ο αλγόριθμος της προσαρμοστικής στοχαστικής απόδοσης καθόδου υλοποιεί τον μηχανισμό του προσαρμοστικού βήματος.

2.3.3 Αγχίγραμμος μετασχηματισμός

Ο μετασχηματισμός που θα χρησιμοποιηθεί για την διαδικασία της κατάτμησης καθορίζει το τι θεωρείται επιτυχία. Ο μετασχηματισμός αντικατοπτρίζει τον επιθυμητό τύπο μετασχηματισμού και περιορίζει τον χώρο λύσης σε αυτόν τον τύπο παραμόρφωσης.

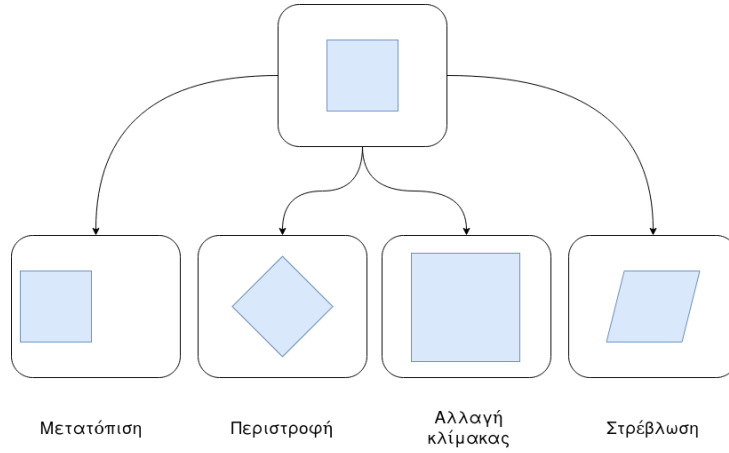
Ο βαθμός ελευθερίας του μετασχηματισμού είναι ίσος με τον αριθμό των παραμέτρων του. Με άλλα λόγια, ο βαθμός ελευθερίας του μετασχηματισμού είναι ίσος με τη διάσταση του χώρου αναζήτησης. Επίσης είναι ανάλογος του με την υπολογιστική πολυπλοκότητα του προβλήματος βελτιστοποίησης.

Ο αγχίγραμμος μετασχηματισμός (Affine) είναι ένας γεωμετρικός μετασχηματισμός που διατηρεί τις γραμμές και τον παραλληλισμό (αλλά όχι απαραίτητα αποστάσεις και γωνίες). Οι ιδιότητες του μετασχηματισμού είναι:

1. Διατηρεί τις ευθείες και τον λόγο αυτών.
2. Διατηρεί τον Παραλληλισμό: δύο ή περισσότερες ευθείες που είναι παράλληλες μεταξύ τους, συνεχίζουν να είναι παράλληλες μετά τον μετασχηματισμό.

3. Διατηρεί ισομετρίες μέσω μεταφορών.
4. Διατηρεί την κυρτότητα συνόλων: ένα κυρτό σετ συνεχίζει να είναι κυρτό μετά τον μετασχηματισμό. Επιπλέον, τα ακραία σημεία του αρχικού σετ αντιστοιχίζονται στα ακραία σημεία του μετασχηματισμένου σετ.

Ο αγγίγραμμος μετασχηματισμός πραγματοποιεί μετατόπιση, περιστροφή, αλλαγή κλίμακας και στρέβλωση.



Σχήμα 1: Μετατόπιση, περιστροφή, αλλαγή κλίμακας και στρέβλωση ενός τετραγώνου.

Έστω $F(x) : \Omega_F \subset \mathbb{R}^D \mapsto \mathbb{R}$ η σταθερή εικόνα, $M(x) : \Omega_M \subset \mathbb{R}^D \mapsto \mathbb{R}$ η κινούμενη εικόνα και $T_\mu(x) : \Omega_F \times \mathbb{R}^P \mapsto \Omega_M$ παραμετροποιημένος μετασχηματισμός συντεταγμένων, όπου $\mu \in \mathbb{R}^P$ αναπαριστά το διάνυσμα των παραμέτρων του μετασχηματισμού. Επίσης P είναι ο βαθμός ελευθερίας του μετασχηματισμού.

Ο αγγίγραμμος μετασχηματισμός ορίζεται ως:

$$T_\mu(x) = A(x - c) + t + c$$

Όπου t είναι το διάνυσμα μετατόπισης και c το κέντρο περιστροφής. Επίσης, ο πίνακας A πραγματοποιεί την περιστροφή, αλλαγή κλίμακας και στρέβλωση. Το διάνυσμα παραμέτρων μ θα αποτελείται από τα στοιχεία a_{ij} του πίνακα A και το διάνυσμα μετατόπισης t . Στις δύο διαστάσεις ο μετασχηματισμός έχει έξι βαθμούς ελευθερίας, ενώ στις τρεις έχει δώδεκα.

2.3.4 Γραμμική παρεμβολή

Κατά την διαδικασία του υπολογισμού της μετρικής αξιολόγησης της κατάτμησης (όπως στη 2.3.1), συγκρίνονται οι τιμές των αντίστοιχων σημείων της σταθερής εικόνας και των μετασχηματισμένων της κινούμενης. Αυτά τα μετασχηματισμένα σημεία της κινούμενης εικόνας μπορεί να μην ανήκουν στο πλέγμα της εικόνας. Γι' αυτόν τον λόγο, είναι απαραίτητη η παρεμβολή των σημείων αυτών.

Η γραμμική παρεμβολή υπολογίζει τον σταθμισμένο μέσο όρο των γειτονικών εικονοστοιχείων, χρησιμοποιώντας την απόσταση ως το βάρος.

2.3.5 Χώρος κλίμακας Gauss

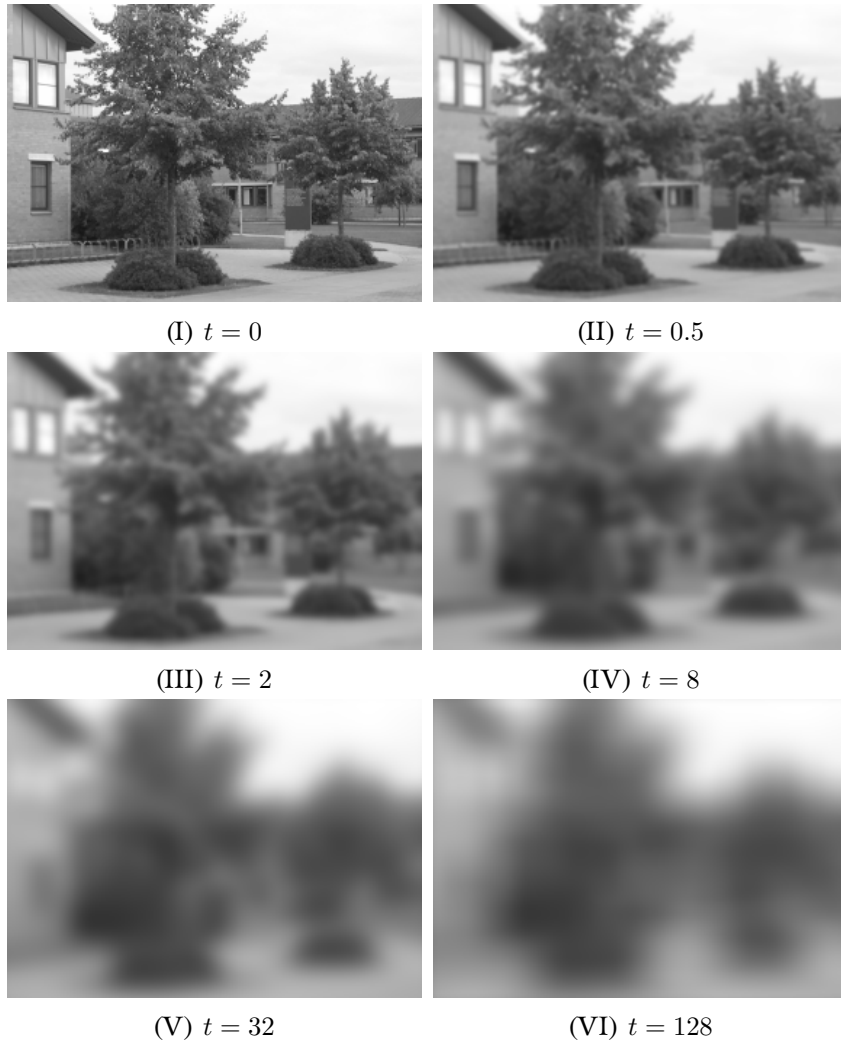
Ένα πρόβλημα που συναντάται όσο αυξάνεται ο χώρος αναζήτησης της κατάτμησης, είναι ότι υπάρχουν μεγαλύτερες πιθανότητες να καταλήξει η βελτιστοποίηση σε κάποιο τοπικό ελάχιστο, με αποτέλεσμα να μην είναι καλό το αποτέλεσμα της. Ένας τρόπος να αυξηθεί η πιθανότητα να βρεθεί το ολικό ελάχιστο είναι να χρησιμοποιηθεί μία ιεραρχική διαδικασία κατά την οποία η πληροφορία των εικόνων θα αυξάνεται από τα αρχικά προς τα τελικά στάδια.

Ο χώρος κλίμακας Gauss είναι ένας τρόπος να επιτευχθεί το παραπάνω αποτέλεσμα [6]. Αυτή η μέθοδος χρησιμοποιεί τον πυρήνα Gauss για να μειώσει την πληροφορία μίας εικόνας. Ο πυρήνας Gauss ορίζεται για τρεις διαστάσεις ως:

$$g(x, y, z, t) = \frac{1}{2\pi t} e^{-\frac{(x^2+y^2+z^2)}{t}}$$

Όπου $\sqrt{2t}$ είναι η τυπική απόκλιση του πυρήνα και t το επίπεδο του χώρου κλίμακας. Ο πυρήνας αυτός συνελίσσεται με μία εικόνα ώστε να μειωθεί η πληροφορία της. Αυτό γίνεται σε πολλαπλά ιεραρχικά στάδια, στα οποία αρχικά η τιμή του t είναι μεγάλη και μειώνεται σε κάθε στάδιο. Ιδανικά με αυτήν την μέθοδο, οι εικόνες θα έχουν ένα ελάχιστο στο αρχικό ιεραρχικό επίπεδο, το οποίο θα είναι κοντά στο επιθυμητό ολικό ελάχιστο. Έπειτα σε κάθε στάδιο αφού αυξάνεται η πληροφορία, το ελάχιστο ιδανικά θα τείνει όλο και πιο κοντά στο ολικό.

Στο Σχήμα 2 παρουσιάζεται η ίδια εικόνα για διάφορες τιμές του t . Από το σχήμα αυτό, παρατηρείται ότι όσο αυξάνεται η τιμή του t τόσο μειώνεται και η πληροφορία στην εικόνα.



Σχήμα 2: Χώρος κλίμακας Gauss για διάφορες τιμές του t .

2.3.6 Δειγματοληψία συντεταγμένων εικόνας

Κατά την διαδικασία της αξιολόγησης μπορεί να χρησιμοποιηθεί η διαδικασία της δειγματοληψίας συντεταγμένων εικόνας ώστε να μειωθεί ο υπολογιστικός χρόνος που χρειάζεται για την αξιολόγηση. Στην διαδικασία αυτή επιλέγονται τυχαίες συντεταγμένες (σημεία στον χώρο της

εικόνας) με ομοιόμορφη κατανομή, ώστε να χρησιμοποιηθούν για την αξιολόγηση. Μέσω παρεμβολής των συντεταγμένων αυτών δημιουργείται η νέα δειγματοληπτική εικόνα.

2.3.7 Μάσκα εικόνας

Αν μία εικόνα (είτε σταθερή, είτε κινούμενη) έχει περιοχές στις οποίες δεν υπάρχουν χαρακτηριστικά ενδιαφέροντος, τότε μπορεί να χρησιμοποιηθεί μία μάσκα εικόνας, ώστε μην να συμπεριληφθούν οι περιοχές αυτές στην διαδικασία της κατάτμησης. Η μάσκα εικόνας είναι μία δυαδική εικόνα, η οποία υποδεικνύει εάν ένα εικονοστοιχείο θα συμπεριληφθεί στην διαδικασία της κατάτμησης. Το μέγεθος της μάσκας είναι ίδιο με το μέγεθος της εικόνας που θα εφαρμοστεί, ώστε να υπάρχει αντιστοιχία μεταξύ των εικονοστοιχείων της εικόνας και των εικονοστοιχείων της μάσκας. Επομένως, κάθε εικονοστοιχείο της μάσκας δηλώνει εάν το αντίστοιχο εικονοστοιχείο της εικόνας θα χρησιμοποιηθεί στην κατάτμηση.

2.4 Μέτρα αξιολόγησης

Τα μέτρα αξιολόγησης χρησιμοποιούνται για να αξιολογήσουν το αποτέλεσμα της κατάτμησης. Συγκρίνουν το αποτέλεσμα της κατάτμησης με το επιθυμητό αποτέλεσμα, ούτως ώστε να παράξουν μία αριθμητική αναπαράσταση της ποιότητας της κατάτμησης. Με αυτόν το τρόπο μπορούν να συγκριθούν διαφορετικοί αλγόριθμοι κατάτμησης και διαφορετικές τιμές για τον ίδιο αλγόριθμο.

Η επιλογή του μέτρου ή των μέτρων αξιολόγησης έχει μεγάλη σημασία για την δημιουργία και τον συντονισμό (επιλογή των κατάλληλων μεταβλητών) του μοντέλου επειδή, είναι το μέτρο που θα χρησιμοποιηθεί για την επιλογή του καταλληλότερου μοντέλου. Η επιλογή αυτή ταυτίζεται με την επιθυμητή συμπεριφορά του τελικού μοντέλου. Για παράδειγμα, αν το επιθυμητό μοντέλο πρέπει να έχει όσο είναι δυνατόν πιο λίγα αληθή αρνητικά, τότε πρέπει να επιλεγεί και το ανάλογο μέτρο αξιολόγησης που να ικανοποιεί τον περιορισμό αυτό.

2.4.1 Συντελεστής ομοιότητας Dice

Ο συντελεστής ομοιότητας Dice είναι μία μέτρηση που εκτιμά την ομοιότητα μεταξύ δύο συνόλων [7]. Είναι γνωστός επίσης και ως δείκτης Sørensen–Dice, συντελεστής Dice και F1 σκορ. Ο συντελεστής ομοιότητας Dice ορίζεται ως:

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|}$$

Όπου X και Y είναι σύνολα, $|x|$ είναι ο αριθμός των στοιχείων του συνόλου x και \cap δηλώνει την τομή δύο συνόλων.

Στην αξιολόγηση της κατάτμησης εικόνας τα παραπάνω σύνολα αποτελούνται από το σύνολο των εικονοστοιχείων που ανήκουν σε μία κλάση. Επομένως ο συντελεστής αυτός υπολογίζει το ποσοστό των σημείων μίας κλάσης που επικαλύπτονται. Η μέγιστη τιμή του συντελεστή είναι 1 και η ελάχιστη 0.

2.4.2 Δείκτης δομικής ομοιότητας

Ο δείκτης δομικής ομοιότητας [8] (Structural SIMilarity index, SSIM) είναι μία μετρική της ομοιότητας δύο εικόνων. Χρησιμοποιεί τις σύγκρισεις της φωτεινότητας, της αντίθεσης και της δομής των εικόνων. Κατά την δημιουργία μίας εικόνας η φωτεινότητα των αντικειμένων της εικόνας εξαρτάται από τον φωτισμό και την ανακλαστικότητα των αντικειμένων αυτών, αλλά και από την δομή τους. Γι' αυτό τον λόγο είναι επιθυμητό να διαχωριστεί η επίδραση της φωτεινότητας έτσι ώστε η δομή να είναι ανεξάρτητη από την φωτεινότητα. Επομένως μπορεί να οριστεί η δομική πληροφορία της εικόνας ως τα χαρακτηριστικά των αντικειμένων της που είναι ανεξάρτητα από την φωτεινότητα και την αντίθεση.

Η φωτεινότητα ενός σήματος ορίζεται ως:

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i \quad (5)$$

Όπου x είναι το σήμα και N ο αριθμός των στοιχείων του. Αν οι τιμές των εικονοστοιχείων μίας εικόνας τοποθετηθούν σε ένα διάνυσμα

τότε μπορεί να χρησιμοποιηθεί ο τύπος (5) για τον υπολογισμό της φωτεινότητας αλλά και των παρακάτω τύπων. Για να αφαιρεθεί η επίδραση της φωτεινότητας στο σήμα θα πρέπει να ισχύει:

$$\sum_{i=1}^N x_i = 0$$

Η αντίθεση της σήματος-εικόνας εκτιμάτε με την διακύμανση του σήματος:

$$\sigma_x = \sqrt{\left(\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right)}$$

Για να αφαιρεθεί και η επίδραση της αντίθεσης θα πρέπει να κανονικοποιηθεί το σήμα, δηλαδή να διαιρεθεί με την διακύμανση του ώστε να έχει μοναδιαία διακύμανση. Επομένως

$$\hat{x} = \frac{x - \mu_x}{\sigma_x}$$

Όπου \hat{x} είναι το κανονικοποιημένο σήμα. Η σύγκριση της φωτεινότητας δύο εικόνων x και y ορίζεται ως:

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (6)$$

Η σταθερά C_1 μπορεί να χρησιμοποιηθεί για να αποφευχθεί υπολογιστική αστάθεια για μικρές τιμές του παρανομαστή. Η σύγκριση της αντίθεσης ορίζεται ως:

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (7)$$

Όπου η σταθερά C_2 χρησιμοποιείται επίσης για την υπολογιστική σταθερότητα. Επειδή η συσχέτιση των κανονικοποιημένων σημάτων \hat{x} και \hat{y} είναι ίδια με την συσχέτιση των μη κανονικοποιημένων σημάτων, μπορεί να οριστεί η σύγκριση της δομής (που είναι ανεξάρτητη από την φωτεινότητα και την αντίθεση) ως:

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3} \quad (8)$$

Όπου η σταθερά C_3 χρησιμοποιείται επίσης για την υπολογιστική σταθερότητα και σ_{xy} είναι η συνδιακύμανση των σημάτων και ορίζεται ως:

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y)$$

Τέλος, χρησιμοποιώντας τις συγκρίσεις της φωτεινότητας, της αντίθεσης και της δομής ((6), (7) και (8) αντίστοιχα), ο τύπος του δείκτη δομικής ομοιότητας ορίζεται ως:

$$SSIM(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [l(x, y)]^\gamma$$

Οι τιμές α , β και γ χρησιμοποιούνται για να ορίσουν την επίδραση κάθε σύγκρισης.

Αναφορές

- [1] D. Zhang, Q. Guo, G. Wu και D. Shen, *Sparse Patch-Based Label Fusion for Multi-Atlas Segmentation*, 2012.
- [2] R. Tibshirani, "Regression Shrinkage and Selection Via the Lasso," *JOURNAL OF THE ROYAL STATISTICAL SOCIETY, SERIES B*, τόμ. 58, σσ. 267–288, 1994.
- [3] L. G. Brown, "A Survey of Image Registration Techniques," *ACM Computing Surveys*, τόμ. 24, σσ. 325–376, 1992.
- [4] B. Zitová και J. Flusser, "Image registration methods: a survey," *IMAGE AND VISION COMPUTING*, τόμ. 21, σσ. 977–1000, 2003.
- [5] S. Klein, J. P. W. Pluim, M. Staring και M. A. Viergever, "Adaptive stochastic gradient descent optimisation for image registration.," *International Journal of Computer Vision*, σσ. 227–239, 2009.
- [6] H. Lester και S. Arridge, "A survey of hierarchical non-linear medical image registration," *Pattern Recognit.*, τόμ. 32, σσ. 129–149, 1999.
- [7] L. R. Dice, "Measures of the Amount of Ecologic Association Between Species," *Ecology*, τόμ. 26, αρθμ. 3, σσ. 297–302, 1945. doi: 10.2307/1932409.
- [8] Z. Wang, A. Bovik, H. R. Sheikh και E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, τόμ. 13, σσ. 600–612, 2004.